The roots of ethnolects

A sociophonological study in Amsterdam and Nijmegen

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration by Linda van Meel.

ISBN: 978-94-6093-203-8
NUR 616

The roots of ethnolects

A sociophonological study in Amsterdam and Nijmegen

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus,

volgens besluit van het college van decanen

in het openbaar te verdedigen op dinsdag 15 maart 2016

om 12.30 uur precies

door

Linda Martina Catharina van Meel

geboren op 26 oktober 1984

te Rotterdam

Promotoren:

Prof. dr. R.W.N.M. van Hout

Prof. dr. F.L.M.P. Hinskens (Vrije Universiteit Amsterdam)

Manuscriptcommissie:

Prof. dr. H.M.G.M. Jacobs

Prof. dr. G.R. Guy (New York University, Verenigde Staten)

Prof. dr. J.A. Villena Ponsoda (Universidad de Málaga, Spanje)

Contents

Dankwoord (Acknowledgements)

Zoals vele promovendi voor mij kunnen beamen, is de route naar een proefschrift er niet een die je alleen aflegt. Ook ik heb de nodige support gehad (zowel direct als indirect) gedurende de verschillende etappes op mijn route. Met dit dankwoord wil ik dan ook iedereen bedanken die mij op wat voor manier dan ook heeft geholpen tijdens de afgelopen jaren.

Op de eerste plaats wil ik de vaste begeleiders op mijn route bedanken, mijn promotoren: Frans Hinskens op het Meertens Instituut en Roeland van Hout op de Radboud Universiteit. Beiden hebben een eigen focus op onderzoek, waarbij die van Frans vooral bij talige verklaringen ligt en die van Roeland bij statistische benaderingen en sociale effecten.

Ook Pieter Muysken heeft vanaf het begin geholpen waar nodig. In feite zou dit proefschrift er zonder Pieter en ook zonder Ineke van der Craats mogelijk helemaal niet zijn geweest. In maart 2009 kreeg ik namelijk een mailtje van Ineke met de vraag of ik even contact op wilde nemen met Pieter. Een kleine twee weken later begon ik binnen het project 'Roots of Ethnolects' (RoE) als onderzoeksassistente. Ineke en Pieter, bedankt dat jullie aan mij gedacht hebben. Tijdens mijn onderzoeks-assistentschap heb ik me vooral bezig gehouden met het maken van de elicitatietaken. Hoewel deze periode geen directe invloed heeft gehad op dit proefschrift (dat zich met de spontane spraakdata bezighoudt), is het toch een belangrijk pad op mijn route geweest. Graag wil ik Pieter, Frans, Roeland, Ariën, Loes en Joop bedanken voor de fijne samenwerking tijdens deze periode.

Na een aantal etappes op het onderzoeksassistenpad, bleek in januari 2010 het volgende deel van mijn route zich op de promovendiweg te bevinden. Ariën van Wijngaarden heeft mij ingewijd in de database met de spontane spraakdata. Bijna alle opnames waren op dat moment al opgenomen door Wouter Kusters, Esther van Krieken, Mohamed Ajaoud, Ariën van Wijngaarden, Hanke van Buren en Özlem Catak. De laatste opnames zijn opgenomen door August de Feniks. Hanke van Buren had al voorbereidend werk voor het onderzoek naar het foneem /z/ gedaan, en er waren al metingen uitgevoerd door David van der Vloed. Sander van der Harst heeft diverse metingen en fonetische analyses op de vier fonemen gedaan en daarnaast een deel van de /a:/ en /ɑ/ gecodeerd. Verschillende stagiaires hebben een deelonderzoek binnen het RoE-project uitgevoerd, waaronder onderzoek naar *hun* als onderwerp en /ɛ/ in 'heb je/ik' dat gebruikt is in het co-variatie-onderzoek. Voor het laatstgenoemde onderzoek is ook data van Ariën van Wijngaarden gebruikt. Dank jullie wel voor jullie bijdragen.

Ook wil ik iedereen bedanken die mij op talig gebied heeft geholpen, ofwel door het geven van informatie of vertalingen, ofwel door beantwoorden van mijn talige vragen. In het bijzonder Abderrahman El Aissati, Janine, Loretta, Josh, Bob en Vera.

Ik ben ook dank verschuldigd aan de diverse (anonieme) *reviewers* voor hun bruikbare commentaar bij de artikelen. *Also, I would like to thank the manuscript-committee, Haike Jacobs, Gregory Guy, and Juan Villena Ponsoda, for their time to read and evaluate my dissertation. A special thanks is provided to Gregory Guy for his useful questions and suggestions.*

De Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), het Meertens Instituut en de Radboud Universiteit dank ik voor hun (financiële) ondersteuning.

Er zijn ook mensen die indirect hebben bijgedragen aan dit proefschrift, die zorgden voor een goede sfeer, een luisterend oor of hernieuwde inspiratie.

I would like to thank all the people I spoke with and listened to at the symposia and workshops I attended, for inspiring me. Ook de bezoekers van mijn 'lab' op het Drongo-festival hebben voor nieuwe inspiratie gezorgd.

Ik heb vele kamergenoten gehad, zowel op de Radboud Universiteit als op het Meertens Instituut, die zorgden voor een prettige werksfeer en waarbij ik terecht kon voor een praatje, een vraag of een luisterend oor: Olga, Bob, Benedicte, Neele, Luis Miguel, Marine, Cefas, Pablo, Ariën, Marieke, Heimir, Anna, Etske, Lotte en Nina. Ook kon ik terecht bij collega's Loretta, Sophie, Robbert, Josh, Margot, Gerrit Jan, Hülya, Kofi, Suzanne, Francesca, alle anderen van de Languages in Contact-groep, Helens groep op de 9e etage, Hella, collega's bij de 'koffie'-pauzes op de 8e etage, al mijn lunchmaatjes zowel op de RU als op 't Meertens, en bij Sandra & Annemarie.

Als laatste wil ik mijn ouders, broer en zus bedanken voor hun liefde, steun, luisterend oor en het oplossen van computerproblemen.

1 Introduction

1.1 Ethnolects

Linguistic variation can result from language contact or from inherent tendencies in language. It is generally known about language contact when it occurred, in contrast to dialect variation, which is always potentially present. These two sources of linguistic heterogeneity have led to two fields of linguistic research, which have developed more or less separately. One branch, language contact studies (e.g. Weinreich 1953) focuses on code-switching and borrowing, as well as the study of pidgins, creoles, mixed languages, and contact-induced language change. The other branch, quantitative sociolinguistic research, investigates inherent variation (cf. Labov 1966; Weinreich, Labov & Herzog 1968), as well as different types of dialects.

There is also a type of language variation where the two fields intersect: ethnolectal variation. In research on bilingualism and language contact, ethnolects are a new topic. Because ethnolectal variation tends to be gradual and continuous, its study requires approaches from quantitative sociolinguistics.

Muysken (2013) makes clear that developments with respect to new ethnolects are often strongly intertwined with and influenced by youth languages and the way they develop. Ethnolects are relatively stable, while youth languages keep changing and can be qualified as highly dynamic. The linguistic features of ethnolects largely pertain to the phonologic and (morpho-)syntactic components, while most defining features of youth languages have a lexical or pragmatic nature. Speakers of youth language are usually conscious of the way they speak, while this holds less for speakers of ethnolects. In the definition of ethnolects, ethnicity is seen as something inherent and self-evident, while ethnicity is a dynamic and evolving property for speakers of youth language. Even more, in youth languages, expressing generational differences may be more important than expressing ethnic boundaries. Nevertheless, the demarcation between ethnolects and youth languages is not always fully transparent and clear-cut. As pointed out by Muysken (2013:742), it seems more likely that the distinctive features of both types can be situated somewhere on a continuum; on one extreme, features are relatively stable, ethnolectal and competence related, while the features on the other extreme are relatively dynamic, youth-language related and performance related.

An ethnolect or some of its features may be shared by different ethnic groups. When a cluster of features is used by speakers of various ethnic groups, it is often called ‘multi-ethnolect’. The term ‘ethnolect’ is then restricted to a variety of which its

speakers typically represent only one specific minority group (Quist 2008:44). Where an ethnolect delimits a specific single ethnic group (e.g. Turkish), a multi-ethnolect will often set apart 'native' and 'immigrant' speakers (Muysken 2013:743). In the words of Clyne (2000:87), the latter "may be termed 'multi-ethnolect' because several minority groups use it collectively to express their minority status and/or as a reaction to that status to upgrade it." An 'ethnolect' may contain both exogenous features coming from other languages and indigenous nonstandard features. Multi-ethnolect features can also result from stabilized second language acquisition phenomena that are not necessary specific to any ethnic group.

Ethnolects may function on a linguistic continuum from non-standard to standard, creating the opportunity for style shifting. Quantitative variationist sociolinguistics interested in style-shifting examine how often certain features are used in different styles or registers, e.g. the proportion of *–in(g)* in casual and in formal contexts (cf. Labov 1972a), by speakers of different age groups, socio-economic backgrounds, cultural groups, etcetera. To mention an example from the Dutch context, Charry (1983) mentions that bilabial (w), a typical ethnolectal Surinamese Dutch feature, was much more present in informal than in formal styles in sociolinguistic interviews. A speaker's repertoire may be investigated as consisting of several styles/registers that are being selected more or less automatically, depending on the characteristics of the communicative context.

Ethnographic sociolinguists assign a much more active and constructive role to the speaker in creating style shifts. They focus on stylization phenomena that "are short-lived, and typically have a studiedly artificial and explicit meta pragmatic quality, often in the form of formulaic phrases or hyperbolic intensifications of a specific style or variety" (Jaspers 2006:134). Stylizations involve a (semi-)conscious representation of identity. The difference between style and stylization is not clear-cut. They may represent the two extremes on a scale from routine to performance with Labovian semi-automatized style shifting on one side and stylization on the other (Jaspers 2006:134). Some ethnolect speakers will take part in stylization "in specific interaction where the projection of their identity is at play" (Muysken 2013:743).

1.2 Ethnolects of Dutch

Ethnolects of Dutch have emerged in two different settings: (post-)colonial and immigrant (cf. Hinskens 2004; Muysken 2010, 2013). Muysken (2013) gives an overview of seven Dutch ethnolects of the past and the present. Four ethnolects are post-colonial varieties. Indonesian Dutch came to rise in the colonial society of the Dutch East Indies from 1900 onwards, varying from a mixed language to a slightly

accented variety. Around the same period, Surinamese Dutch arose, in which much local vocabulary was added to the (phonetically 'colored') Dutch as spoken in Paramaribo. Specific grammatical and semantic characteristics came in when traditionally creole or Sarnami Hindustani speaking populations in the colony started learning Dutch as a second language. Moluccan Malay Dutch and Antillean Dutch emerged a few decades later (around 1920 and 1950 respectively), developing originally outside the Netherlands as well, but brought to the Netherlands during immigration.

Three ethnolects have their origin in the Netherlands, as a consequence of the arrival of groups of new immigrants. Yiddish/Jewish Dutch is the oldest documented ethnolect (from about 1750 onwards) and came to rise when groups of Ashkenazic Jews came to the Netherlands and started to learn and speak Dutch. Two larger groups of labor immigrants came to the Netherlands much more recently, from the 1970s onwards: immigrants from Morocco and from Turkey. Their ethnolects, in particular that of the Moroccans, have been studied in the last two decades in a series of studies (Cornips 2002, 2008; Dorleijn & Nortier 2006; Hinskens 2011, 2014; Nortier 2008; Nortier & Dorleijn 2008; van Krieken 2004, 2005). Muysken (2013) gives an overview of some of the variable features of Moroccan and Turkish ethnolects. These include: reduction or omission of final -t; realization of /r/, /z/, /y/, plosives, and the sequence /sx/; realization of vowels /ɛ/, /ɛi/, /œy/; staccato speech style; gender reduction; omission of indefinite article; *er*-omission; omission of subject or object pronouns; non-standard preposition selection; special forms of negation; indication of possessive relations; and to a limited extent lexical borrowings from traditional language. However, most of these features are incidental observations and need to be investigated (more) systematically and in linguistic depth.

1.3 State of affairs in international research

This section gives a broad overview of ethnolectal studies in several European countries and countries outside of Europe. Given the abundance of studies, the overview cannot be exhaustive.

Most research on recent immigrant ethnolects in Europe is concentrated in Western Europe. Several research projects were carried out in the Scandinavian countries. Ethnolectal research in this part of Europe started with the study on a variety of Swedish spoken in the suburbs of Stockholm where many immigrants and their descents live. The variety was baptized Rinkeby Swedish after the first immigrant suburb studied (Kotsinas 1988, 1998; Stroud 2004). The largest immigrant group in Rinkeby is the one of Turkish origin (16%), followed by the Greek, Finnish, and Iranian groups, and 60 smaller immigrant groups. Of all inhabitants in Rinkeby

only 26% are “native Swedes” (Kotsinas 1998), i.e. without immigrant backgrounds. With that many different groups and even more heritage languages, Rinkeby is a multi-lingual suburb. Several multi-lingual suburban areas were studied in the project “Language and language use among young people in multi-lingual urban settings” (cf. Källström & Lindberg 2011), which investigated ethnolectal varieties and their use in the three major cities of Sweden (Stockholm, Gothenburg and Malmö). The national project “Linguistic Development in Urban Environments” in Norway tried to identify new varieties of Norwegian that are influenced by multi-lingual environments. The main focus was on multi-ethnolectal linguistic practices in Oslo (e.g. Svendsen & Røyneland 2008). Also in Denmark, multi-ethnolects were studied. Quist (2000, 2008) examined multi-ethnolect from second immigrant youth with different backgrounds in Copenhagen. The heritage languages of her participants were Danish, Berber, Turkish, Kurdish, Palestinian Arabic, Urdu, and Serbian. Jørgensen (2005) studied the range of varieties used by “Turkish-Danish grade school students”.

In the United Kingdom, the focus is on varieties that emerged in post-colonial settings, i.e. varieties of people of South Asian and of Afro-Caribbean descent (e.g. Hewitt 1986; Rampton 1995, 1998, 2011). Instead of ‘British Asian’, Harris (2006) prefers the term ‘Brasians’ in her study in West London, while Stuart-Smith, Timmins & Alam (2011) alternate between the terms ‘Glasgow Asian’ and ‘Glaswasian’ in their study in Glasgow. Most studies on ethnolects in the UK were situated in and around London. One of the varieties studied is 'Multi-cultural London English' (MLE) (cf. Cheshire et al. 2008, 2011; Fox, Khan & Torgersen 2011; Kerswill, Torgersen & Fox 2008). The Birmingham study (Khan 2006) explores the extent of inter-ethnic linguistic variation among adolescents in Birmingham (see also Fox, Khan & Torgersen 2011).

Immigrants, formerly called *Gastarbeiter*, lit. ‘guest worker’, with Turkish backgrounds represent the largest and most visible immigrant group in Germany (Kern 2015). Several labels are used to describe the variety of young second and third generation immigrant (Turkish) German: e.g. ‘Türkendeutsch’ (‘the Turks’ German’), ‘Türkenslang’ (‘the Turks’ slang’), and ‘Türkischdeutsch’ (‘Turkish German’). Another derogatory term is ‘Kanak Sprak’ which has xenophobic associations. “Semantically, it supports an *othering* of (a) the speakers themselves as foreign and (b) of their way of speaking as a different language, ‘Sprak’” (Wiese 2013:6). ‘Kanak Sprak’ can be used for stylized practices (e.g. Deppermann 2007), but Wiese (2009, 2013) prefers the term ‘Kiezdeutsch’ for the non-stylized variant; in her view, ‘Kiezdeutsch’ does not

carry the pejorative connotations of 'Kanak Sprak'. It makes clear that it is a variety of German ('deutsch'). Rather than focusing on the largest group of speakers, i.e. the Turkish Germans, the term 'Kiezdeutsch' focuses on the location where it is spoken, namely, in the *Kiez*, the (neighbor-)hood, "without implying any restrictions, say, to particular ethnicities or migrant vs. non-migrant backgrounds of its speakers" (Wiese 2013:6).

Research on German ethnolects centers in Berlin (e.g. Freywald et al. 2011; Kern 2011a; Queen 2012; Selting 2011; Selting & Kern 2009; Şimşek 2011; Wiese 2009, 2013), Mannheim (e.g. Kallmeyer & Keim 2003; Keim 2002; Keim & Knöbl 2011), and Hamburg (e.g. Dirim & Auer 2004).

In Belgium, research is done to examine style practices by people of immigrant Moroccan descent (Jaspers 2006, 2008) in Antwerp. Other projects investigated style practices by people of both immigrant and non-immigrant descent who use a variety (called *Citétaal*) which is 'a melting pot language, based on Dutch but with a high amount of code mixture from immigrant languages, mostly Italian and Turkish' (Marzo & Ceuleers 2011). *Citétaal* is rooted in local and regional coalminers' jargon.

The first set of studies in France focused on Pied Noir French / Algerian French and described borrowings from other languages like Arabic and Spanish (e.g. Mazzella 2005; Meunier 2001). A second set of ethnolectal studies examined intonation contours (Fagyal 2005; Fagyal & Stewart 2011) and variation in the realization of certain consonants (Jamin 2004, 2005; Jamin, Trimaille & Gasquet-Cyrus 2006) in the multi-ethnic neighborhood (*banlieue*) La Courneuve of Paris as well as in multi-ethnic neighborhoods in Grenoble and Marseille.

In North America, ethnic varieties have been studied in the USA as well as in Canada. This chapter will only mention a few of these studies. The most well-known is that of Labov in New York City (Labov 1966, 1972b) on African American Vernacular English. This variety has been investigated in many studies ever since in a range of local (typically urban) situations across the country. Another group studied is the English-Spanish bilingual community. For example, Ma & Herasimchuk (1971, 1972) studied the variety of Puerto Rican speakers in Jersey City, Chicanos in the Southwest are involved in the study of Bills (1977), and Eckert's (2008a) study concerned Latino Americans in Northern California. Also other immigrant based varieties are investigated, e.g. Carlock and Wölck (1981) examined how citizens of Buffalo, upstate New York, identify fragments of English by third generation of Italian, Polish and German descent. They found that the younger the raters are the more

they identify the fragments as belonging to someone from a particular neighborhood instead of a specific ethnic group. While earlier mentioned studies in the USA examined certain features that characterize a given variety, Cutler (2008) studied which stylized language and lifestyle practices teenage immigrants from New York City used to express their affiliation with hip-hop culture.

In Canada, the project "Contact in the City: Ethnolinguistic Variation in Toronto English" was carried out in Toronto (Hoffman & Walker 2010), among others investigating the role of ethnicity. Boberg (2004) examined ethnic patterns in the phonetics of Montreal English by Montrealers of Irish, Italian, and Jewish ethnic origin.

Research on ethnolects has also been carried out at other continents. For instance, in South Africa, where several varieties of English are spoken by members of different ethnic groups, ethnolects have been studied by Mesthrie (2002, 2012). In Australia, Clyne (2003) has done influential research on ethnolectal varieties of Australian English.

For every country specified in the overview in Appendix 1, there is at least one study that goes into phonetic or phonologic variables. Almost all research on ethnolectal varieties of English (in UK and USA) in Appendix 1 includes phonetic or phonologic variables. As of roughly the year 2000, there are more studies that delve into interpretive, speaker-oriented concepts such as stylizations, practices, and stances. Also more studies do not specify the ethnic background of their participants / informants, but use descriptions as 'people in multiethnic and multilingual neighborhoods', native vs. non-native, multi-ethnic vs. mono-ethnic neighborhoods.

In all these research projects and studies, we can distinguish roughly two kind of approaches (cf. Hinskens 2011): an approach that focuses on language and its structural properties: the structural-linguistic approach, and an approach that focuses on language and its use in communicative behavior: the ethnographic approach. The more structural-linguistic approach tries to unravel the laws, generalizations and restrictions on variable features of 'ethnolects', 'multi-ethnolects' or 'multi-cultural varieties'. Researchers following the ethnographic approach prefer to speak about 'styles' or '(pan-) ethnic styles' (see Kern 2011b:9) between which speakers may freely shift depending on the evoked associations with different groups. Some researchers of the latter approach therefore prefer to call this approach the 'stylistic (practice) perspective' (cf. Eckert 2012; Marzo & Ceuleers 2011). Whereas in the structural-linguistic approach variation is seen as reflecting social categories (e.g. gender, ethnicity, and social networks including 'culturally' and 'religiously' defined

networks), in the ethnographic approach variation is seen as both reflecting and constructing social meaning. Ethnographic research is (mainly) qualitative and interpretative, while structural-linguistic research is usually quantitative, often in the Labovian tradition, and focuses on features of linguistic structure, their origin and distribution.

In our dichotomy, 'structural-linguistic' and 'ethnographic' constitute the poles of a continuum. In the great majority of cases, the position of an author can be determined well in relation to the two poles, but extreme positions are rarely taken because most studies reflect briefly on the relevance of the other perspective.

While we outline these two opposite poles, Kern (2015) distinguishes three points of view: 'varionist sociolinguistic studies', 'ethnographic-interactional studies' and 'interactional linguistic studies'. 'Varionist sociolinguistic studies' focus on the formal description of the linguistic features that characterize an ethnic variety in contrast to more standardized varieties. In 'ethnographic-interactional studies' ethnographic information plays an important role in studying language behavior. 'Interactional linguistic studies' "aim to investigate communicative practices in everyday conversation" with an emphasis on how elements are used for the organization of talk-in-interaction. Contrary to the other two kinds of studies, 'interactional linguistic studies' use the term the term 'style' instead of 'variety'. In our dichotomy, studies like the 'ethnographic-interactional studies' and 'interactional linguistic studies' of Kern (2015) occupy positions which can be considered as intermediate in certain respects.

The first studies with a mainly ethnographic perspective in the overview in Appendix 1 started appearing about a decade after the earliest studies from a mainly structural-linguistic perspective. By far most of the ethnographic studies in the overview in Appendix 1 appeared in the first decade of the 21st century, while there is no such a high peak for the structural-linguistic studies. The structural-linguistic perspective is first of all directed towards variation in the sound components, somewhat less towards (morpho-) syntax and relatively little towards lexicon and pragmatics. The ethnographic approach also focuses on variation in the sound components and lexicon, but is more directed towards pragmatics than to structures.

In Europe, and certainly in the Netherlands, much ethnolectal research was either ethnographic and conversational, and thus interpretive, or it was more descriptive in aiming at an inventory of features observed in a certain ethnolect or ethnic style. Systematic research with linguistic depth of analysis was lacking before 2005. To resolve this, a project was set up in the Netherlands to investigate the origin of features of modern Dutch ethnolects systematically: 'The Roots of Ethnolects, an

Experimental Comparative Study'. Section 1.4.1 will outline the aims of the project, followed by a sketch of the research design (Section 1.4.2) and some information about the collected data in the project database (Section 1.4.3). Section 1.5 will give some background information on the demography and language history of the involved cities, Amsterdam and Nijmegen. Section 1.6 will introduce the core corpus, i.e. a selection of data from the project database. The research questions of the thesis are presented in Section 1.7.

1.4 The overarching project: Roots of Ethnolects

1.4.1 Aims

The project 'The Roots of Ethnolects, an Experimental Comparative Study' was founded in 2005 by Pieter Muysken[1] and Frans Hinskens[2]. The study was set up to explore the different sources, the 'roots', of ethnolects, including the interaction between second language acquisition, multi-lingual language use, and ingroup/outgroup dynamics in urban settings. The project aims at investigating the emergence, position and social spread of two young ethnolects of Dutch in Amsterdam and Nijmegen. These are spoken by second-generation migrants of Turkish and Moroccan descent. The project focuses on phonological/phonetic and morpho-syntactic features and encompasses the following set of questions:

(a) Which aspects of language use (components of the grammar) characterize ethnolects as distinct varieties?
(b) To what extent are ethnolects based on local non-standard-varieties?
(c) To what extent are ethnolects based on interference from the original language of the ethnic group in question?
(d) To what extent can we reduce features of ethnolects to properties resulting from processes of language acquisition?
(e) Are ethnolects specific for an individual ethnic group, or do they reflect a more global non-native identity? Are 'uniquely' ethnolectal traits which reflect a more global non-native identity acquisition-driven?
(f) To what extent can speakers of an ethnolect shift to more standard varieties and to non-ethnic non-standard varieties?
(g) Is there any evidence of spread of ethnic varieties to peers outside of the ethnic group? If so, do 'uniquely' ethnolectal traits spread to peers outside the ethnic networks merely because of their 'covert prestige' or also / rather because they represent less marked options?

[1] Radboud University, Nijmegen.
[2] Meertens Institute & VU Amsterdam.

(h) How widely are ethnolects distributed in the country?
(i) To what extent do ethnolects resemble code-switched varieties and contain overt lexical and grammatical material from several languages?

1.4.2 Research design

To address the questions posed in the Roots of ethnolects project, a structured semi-experimental factorial design was set up. The aim was to examine both natural speech and more controlled elicited speech in two cities, Nijmegen and Amsterdam. The targeted informants were 10-12 and 18-20 years old males with 'white' Dutch, Moroccan and Turkish backgrounds. As for the 'white' Dutch group, both speakers with and without inter-ethnic ties had to be recruited. Table 1 presents the original speakers design.

Just labeling our participants *Dutch*, *Turkish* and *Moroccan* will not do justice to the fact that the participants with a Turkish or Moroccan background were also born in the Netherlands. It is hard, if not impossible, to devise proper, politically correct, indisputable labels for our informant groups. We opted for *Turkish-Dutch*, *Moroccan-Dutch* and *'white' Dutch* as group labels. The term 'white' is often used in North-American studies, but is controversial in other parts of the world. As *'white' Dutch* is not always accepted, *endogenous Dutch* is used in some of our chapters.

Table 1 The speakers design for the *Roots of ethnolects* project (Hinskens 2011, 2013)

City	Main language background	Age group	
		10 to 12 years old	18 to 20 years old
Amsterdam	Moroccan	6	6
	Turkish	6	6
	'white' Dutch, strong inter-ethnic ties	6	6
	'white' Dutch, weak inter-ethnic ties	6	6
Nijmegen	Moroccan	6	6
	Turkish	6	6
	'white' Dutch, strong inter-ethnic ties	6	6
	'white' Dutch, weak inter-ethnic ties	6	6

The *natural speech* data had to be collected by having speakers of the same age group talk to each other. All speakers except the 'white' Dutch with weak ties had to have conversations in three settings: with someone from their own background and with someone of each other background. The 'white' Dutch without interethnic ties would only have conversations with other 'white' Dutch without ethnic ties. The

controlled elicited speech data consisted of several specific elicitation tasks to be performed by the speakers.

1.4.3 Database

The natural speech data were collected by Wouter Kusters, Esther van Krieken, Mohamed Ajaoud, Ariën van Wijngaarden, Hanke van Buren, Özlem Catak and August de Feniks. They recorded 155 conversations by 93 speakers. The distribution of the speakers is presented in Table 2. One would expect more conversations from 93 speakers, but unfortunately, not all speakers were recorded in conversations with speakers from each of the three different backgrounds. Meta-data of the speakers were collected by means of a short interview. As the speakers were already facing three to four hour of participating, the interview was kept short. The interview included questions on which languages they speak well and which languages they speak with family / friends, on the background of family and friends, and on their living history. All meta-data topics and interview questions can be found in Appendix 2. Due to unanticipated complications during the fieldwork, only limited information is available about the language skills, linguistic profiles, and family and friend structures of several speakers. Therefore, it is not possible to use an ethnic orientation index like Hoffman & Walker (2010) or linguistic profiles as suggested in Fraurud & Boyd (2006, 2011) as a social factor.

For practical reasons, the controlled elicited speech data was collected in a later stage with other speakers than the ones providing the natural speech data. These data will not be discussed here.

Table 2 The actual number of speakers recorded in the Roots of Ethnolects spontaneous speech database

City	Main language background	Age group	
		10 to 12 years old	18 to 20 years old
Amsterdam	Moroccan	7 *	6
	Turkish	5	6
	'white' Dutch, strong inter-ethnic ties	7 *	4
	'white' Dutch, weak inter-ethnic ties	6	3
Nijmegen	Moroccan	6	7 *
	Turkish	6	6
	'white' Dutch, strong inter-ethnic ties	6	7 **
	'white' Dutch, weak inter-ethnic ties	6	5

* Includes an extra speaker who served only as a conversation partner; this happened five times (A12M0X, A12D0X, N20M0X, N20D0X, N20D0Y)

1.5 Amsterdam and Nijmegen

Figure 1 Locations (in red) of Amsterdam and Nijmegen in the Netherlands

As said above, the research took place in two cities in the Netherlands: Amsterdam and Nijmegen. Their locations are depicted in Figure 1. Amsterdam has been the leading Dutch city in economic, political and cultural respects for centuries. Nijmegen may claim to be the oldest city of the Netherlands with Roman origins (cf. http://english.nijmegen.nl/). The *city* celebrated its *2000th anniversary* in 2005. Due to its location near the Waal River, the main branch of the Rhine River in the Netherlands, Nijmegen has always been important for trade and commerce.

Nijmegen was a fortified town until 1874, when the city was set free from the fortified status and got the opportunity to grow, i.e. to build and expand in the area outside the ramparts.

1.5.1 Demography

Amsterdam has a much larger population than Nijmegen. In 2005, i.e. the start of the project 'The Roots of Ethnolects', Amsterdam had 742,951 inhabitants (O+S Amsterdam 2009), while Nijmegen had 158,215 (O&S Nijmegen 2014).[3] In interpreting these figures, it has to be taken into account that both cities have smaller neighboring municipalities that are oriented towards the two cities, meaning that the urbanized area is in fact larger than the figures indicate.

Figure 2 shows the percentage of people of 'white' Dutch (i.e. *autochtoon*, 'autochthonous'), Turkish and Moroccan origin, respectively, in Amsterdam and Nijmegen. As of 2006, Amsterdam and Nijmegen use the term *herkomstgroepering* 'group of origin', as defined by the national *Centraal Bureau voor de Statistiek (CBS,* 'Central Statistical Office'), instead of the former classification in terms of *ethnicity*.

The group of origin of 'someone born abroad with at least one parent who was born abroad' is the country where he or she is born (CBS 2006a, 2006b). E.g. the origin of someone born in Turkey is Turkish, but the origin of someone living in the Netherland, born in Morocco with parents born in Turkey, would be of Moroccan origin.

The group of origin of 'someone born in the Netherlands who has at least one parent born abroad' is the native country of the mother, unless that is the Netherlands. In that case, the group of origin is determined by the father's native country (CBS 2006a, 2006b). E.g. the origin of someone with a mother born in Turkey and a father born in the Netherlands is Turkish. The origin is also Turkish when the mother's native country is the Netherlands and the father's Turkey. However, the origin will be Moroccan when the mother is born in Morocco and the father in Turkey.

'A person whose parents were both born in the Netherlands' is labelled *autochtoon,*[4] i.e. 'someone with a Dutch background', irrespective of the country the person is born in (CBS 2006a, 2006b). E.g. if someone is born in Turkey, but his/her parents were both born in the Netherlands, the person is not of Turkish origin, but considered an *autochtoon* (having a Dutch background, i.e. a Dutch origin).

[3] Recent numbers from 2014 (O+S Amsterdam 2014; O&S Nijmegen 2014) show that Nijmegen (168,344) is in terms of population still one fifth the size of Amsterdam (811,185).

[4] CBS uses the term *autochtoon* instead of '*Nederlands'* (Dutch) for people with a Dutch background.

Amsterdam turns out to be a more multi-cultural city than Nijmegen, as only about half of the Amsterdam population is of 'white' Dutch origin (i.e. *autochtoon* 'autochthonous'), while this is about 75% in Nijmegen, as can be seen in Figure 2. The figures are fairly stable over the years. Amsterdam has more inhabitants of Turkish and Moroccan origin (both in number and percentages) than Nijmegen. The proportions of people with a Turkish origin and with a Moroccan origin differ in both cities. While there are nearly four percent points more people of Moroccan than Turkish origin in Amsterdam, the picture in Nijmegen is reverse with about 1% more people of Turkish than Moroccan origin, both groups scoring lower than 3.3%.

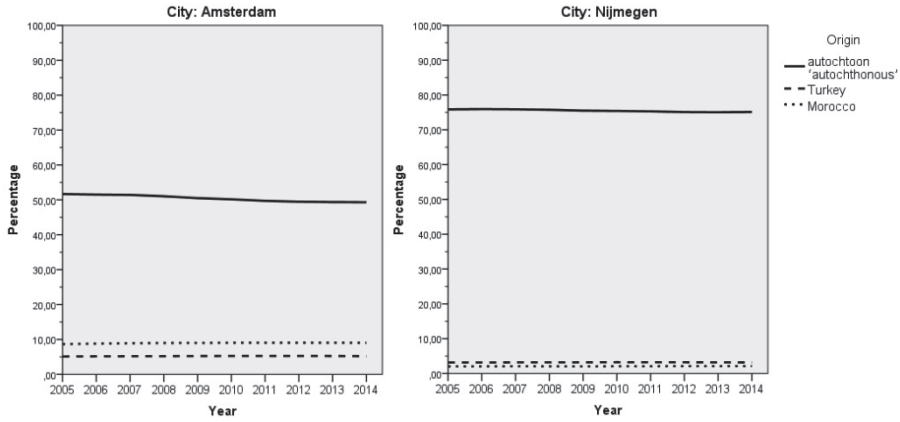


Figure 2 Population 'group of origin', displaying the percentage of people of Dutch origin (i.e. *autochtoon*, 'autochthonous'), people of Turkish and people of Moroccan origin (O+S Amsterdam 2009, 2014; O&S Nijmegen 2014)

As it is hard to find persons in Amsterdam without any contacts with Moroccan-Dutch or Turkish-Dutch people, the 20-year-olds with a Dutch origin without ethnic ties (participant code: A20C) were recruited in Aalsmeer. Aalsmeer, which is located 15 km southwest of Amsterdam, is part of the *Stadsregio Amsterdam*, 'City Region of Amsterdam, i.e. the metropolitan region of Amsterdam', a cooperation of 16 neighboring municipalities (cf. http://www.stadsregioamsterdam.nl/), see Map 2 in Appendix 2. Table 3 presents the number and percentage of people of Dutch origin (i.e. *autochtoon* 'autochthonous'), of Turkish origin and of Moroccan origin in Aalsmeer in the year of recording (2010). Aalsmeer and Amsterdam are part of the same larger dialect area. Aalsmeer clearly has a high percentage of inhabitants with a Dutch origin.

Table 3 Population of AALSMEER; 'group of origin', displaying the number and percentage of 'white' Dutch (i.e. *autochtoon* 'autochthonous'), people of Turkish and of Moroccan origin of the year of recording (CBS 2014; O+S Amsterdam 2011)

background core participant (n)	Year(s) of recording	population, number of residents, *n* (%)	origin: *autochtoon* 'autochthonous', *n* (%)	origin: Turkey, *n* (%)	origin: Morocco, *n* (%)
A20C (3)	2010	29,187 (100)	24,783 (84.91)	341 (1.17)	75 (0.26)

1.5.2 Language history

1.5.2.1 Amsterdam

The dialect of Amsterdam had changed over centuries because of the influx of economical and religious refugees and because of the migration of people from the surrounding regions. Back in medieval times, Amsterdam was a village with a dialect that resembled the surrounding northern dialects (Berns 2002:16; Daan 1948:11). The village developed into a fast growing city, with an enormous flourish around the end of the 16th century. The dialect gradually started to diverge from the surrounding dialects. Berns (2002:19, 20) shows one grammatical and seven phonological Northern Hollandic features (four of which concern vowel quality) which are no longer to be found in the present-day Amsterdam dialect. In 1585, Spain took over the city of Antwerp. Many high educated and rich people from Antwerp fled to Holland, especially to Amsterdam. The refugees from the Antwerp area may well have imported the monophthongization of the diphthongs /ɛi, au, œy/, a characteristic of the present-day Amsterdam dialect, as this phenomenon can also be found in a group of Flemish and Brabant dialects of Dutch (Hinskens & Muysken 2007). Not long after the Antwerp refugees, from 1593 onwards, another group of religious refugees arrived in Amsterdam: the Sephardic Spanish and Portuguese Jews (cf. Hinskens & Muysken 2007). From 1618 onwards, Jewish refugees from Central and Eastern Europe (Ashkenazim) migrated to the Netherlands (cf. Hinskens & Muysken 2007). The present-day Amsterdam dialect as well as more general colloquial Dutch contain many originally Jewish lexical items, i.e. words of Hebrew-Aramaic origin as well as Dutch Yiddish words of Germanic origin (cf. Hinskens & Muysken 2007). Yiddish, the main vernacular of the Ashkenazim, has also had phonetic effects on the Amsterdam dialect. The characteristic Amsterdam 'grave' and slightly palatal pronunciation of /s/ may be derived from the so-called 'ghetto-pronunciation' of Dutch or, indirectly, from Yiddish (Hinskens 2004).

In the 17th century, economic refugees came to Amsterdam. People from the other provinces of the Dutch Republic came looking for labor. The largest group of immigrants from outside the Dutch Republic came from Norway, Sweden, Denmark and Germany. Part of the migration was circular, i.e. with short stays to earn lots of

money in short time, or for education or work experience (Kuijpers 2005:17). During the Thirty Years War (1618-1648) immigrants from Germany fled to the Netherlands. In Amsterdam, they had a low status. Another wave of, this time religious, refugees entering Amsterdam were the Huguenots (a term used for French Protestants in 16th and 17th century) after the revocation of the Edict of Nantes in 1685. Only a century later, in the decades after 1795, French loan words started entering the Dutch vernacular via the language of upper class that was strongly oriented towards French (Daan 1948:8). More French loan words have been preserved in the Amsterdam dialect (as well as other dialects) than in (later) standard Dutch varieties (Hinskens & Muysken 2007). By the end of the eighteenth century, probably 19 different dialects existed in Amsterdam (Daan 1948:12). The dialect varieties differed from each other mainly phonologically (particularly the quality and quantity of the vowels) and lexically (Hinskens & Muysken 2007). Later, when people were willing to move or work in other areas, the differences became smaller. Daan (1948:35) state that by the time of 1948 not much was left of the differences between the various dialects, although long established Amsterdam habitants could still hear the difference between two certain dialect varieties, i.e. *Kattenburgs* and *Jordaans* (1948:35).

1.5.2.2 Nijmegen

The dialect of Nijmegen has become a sociolect, nowadays, with a strong social stratification and an overt negative evaluation (van Hout 1999; van Hout, Roelofs & van Stijn 2006). This is mainly caused by two major historical events. The first event took place at the end of the 19th century. Nijmegen originally was a fortified town and was not allowed to grow outside of the ramparts. When the fortified status was revoked in 1874, growth started involving migration from rural areas to the city. Higher educated people moved from the *benedenstad* (the old town) to other areas of Nijmegen causing the *benedenstad* to fall into social decay, losing its function as city center, and becoming the living area of underprivileged groups. The second event relates to the Second World War, when Nijmegen was severely bombed. The disrepair and redevelopment led the closed community of the *benedenstad* to fall apart. Together with increasing education, this led to the demise of the dialect as such. Van Hout (1999) distinguishes three dialect variants (besides Standard Dutch): Old *Nijmeegs*, New *Nijmeegs* and Newspeak, old *Nijmeegs* being the oldest variant. With the social decay of *benedenstad*, the old dialect fell into disrepute for higher educated people. Because of a tension between the old dialect and the standard language and between old speakers and new speakers (immigrants), new patterns of variation emerged, leading to a new variant of *Nijmeegs*. For example, the old [i]- and [y]-forms disappeared and got a monophthongal pronunciation of the diphthongs of

Standard Dutch, e.g. Old *Nijmeegs* [ki(:)kə], New *Nijmeegs* [kɛ:kə], *kijken* 'to watch'. The fricatives became voiceless, e.g. Old *Nijmeegs* [zɛs], New *Nijmeegs* [sɛs], *zes* 'six'. These dialect features now occur especially in speech of low educated people of Nijmegen.

1.6 The core corpus

Table 4 Overview of the participants of the 'core corpus' indicated by their codes

Background	Moroccan-Dutch		Turkish-Dutch		'white' Dutch			
Inter-ethnic ties?	Yes		Yes		Yes		No	
Years of age	10-12	18-20	10-12	18-20	10-12	18-20	10-12	18-20
Amsterdam	A12M01 A12M02 A12M04	A20M03 A20M04 A20M05 A20M06	A12T01 A12T02 A12T03	A20T03 A20T04 A20T05 A20T06	A12D01 A12D02	A20D04 A20D05 A20D06	A12C01 A12C02 A12C03	A20C01 A20C02 A20C03
Nijmegen	N12M01 N12M02 N12M03	N20M01 N20M04 N20M05	N12T01 N12T02 N12T03	N20T01 N20T04 N20T05	N12D01 N12D02 N12D03	N20D01 N20D04 N20D05	N12C01 N12C02 N12C03 N12C04	N20C01 N20C02 N20C03 N20C05

The research in this thesis deals with *natural, spontaneous speech data*. The data used were selected from the full database, with the aim to optimize the possibilities to trace the different primary factors of the research design. The most realistic aim given the data collected was to go for a minimum of three speakers per table cell (cf. Table 2, 16 cells), for whom all relevant speech data were available and who met the selection criteria. In total 51 speakers were selected as 'core participants'. Table 4 gives an overview of the participants chosen as 'core participants'. While the aim was to have the age groups 10-12 and 18-20, some core participants in Nijmegen were actually a year younger or older. In the 10 to 12 year old age group, one boy (N12C03) was actually nine years old. In the 18-20 years old age group, four participants were 17 years old (N20D05, N20T01, N20T04, N20T05), while three participants were 21 years old (N20C03, N20D01, N20M04).

All participants were born and raised in resp. Amsterdam / Aalsmeer and Nijmegen, except for one Nijmegen speaker. Participant N20M01 was born in Morocco and came to the Netherlands when he was about 6 years old.

One boy classified as 'white' Dutch (N20D04) would be classified as Italian according to the official definitions of CBS (see Section 1.3.2.1), as his mother was born in Italy. However, his mother moved to the Netherland when she was three years old and only speaks Dutch with her son.

More information on the language skills and linguistic profiles of the core participants can be found in Chapter 2, Section 2.4.1.

1.6.1 The schools involved

1.6.1.1 Schools of 10-12 year old participants.

All 10-12 year old participants were recruited in primary schools. The participants with ethnic ties attended so-called 'black' schools. This is "the common, though somewhat stigmatizing designation used in The Netherlands" (Nortier & Dorleijn 2008) for schools with a majority of 'non-Western allochtonous'[5] students. The participants of Dutch origin without ethnic ties were attending schools not marked as being 'black'. The schools in Nijmegen were chosen based on two reports from 2003 with data on pupil composition and size of primary schools in Nijmegen (Commissie Spreiding en Bereikbaarheid Nijmegen 2003; O&S Nijmegen 2003). Table 1 of Appendix 4 gives the primary schools from which our core participants came. It also gives (a) the number of children attending the school in the school year 2001-2002, (b) the percentage of non-Western 'allochtonous' children in the 'base generation'[6], (c) the percentage of 'non-Western allochtonous' children attending the school, and (d) whether the school is considered 'black'.

No report on the ethnic composition of the schools in Amsterdam was available at the time the primary schools were selected. The schools were selected based on the impressions and observations of the project members. In 2008, a study was started by O+S Amsterdam on the composition of the school population. Two of their reports (O+S Amsterdam 2008, 2010) give percentages for the schools from which our participants were selected. These percentages suggest that our impressions were correct. Table 2 of Appendix 4 shows the primary schools in Amsterdam from which our core participants came. It also gives the percentage of children of non-Western origin of school age who live nearby a school (i.e. Amsterdam is split up in 'neighborhood combinations', see Section 1.6.2), the percentage of 'non-Western allochtonous' children attending the school, and whether the school has more than 50% 'non-Western *allochtonen'*.

---

[5] *Non-Western allochtonen*: if a child or one of the parents was born in a non-Western country, the child is counted as *a non-Western allochtoon*. *Non-Western* children are mainly Turkish, Moroccan, Antillean, Aruban and Surinam children, but also (among others) Iranian/Iraqi and Indonesian children. The largest groups are the Turkish and Moroccan children. Note that there are also W*estern* children. These are mainly German, Belgium and British children (Commissie Spreiding en Bereikbaarheid Nijmegen 2003:13).

[6] Base generation: "the number of children from 4 until and including 11 years old and 30% of the 12 year olds, living in the neighborhood of a school" (Commissie Spreiding en Bereikbaarheid Nijmegen 2003:13, our translation).

1.6.1.2 Schools of 18-20 year old participants.

The Dutch educational system has three levels of higher education. From low to high, these are: *MBO* ('secondary vocational education'), HBO ('higher professional education') and WO ('academic higher education'). Based on the intuitions of the project team as well as on the report *Onderwijsmonitor 2003* (O&S Nijmegen 2004), it was decided to recruit our 18-20 year olds at *MBO* level, as it would be easier to find participants with inter-ethnic ties. The *Onderwijsmonitor 2003* (O&S Nijmegen 2004:83) reports that 39% of the 2489 students living in Nijmegen[7] and attending the *MBO* ('secondary vocational education') in school year 2002-2003 are non-Western. When looking at the participation of 17-22 year olds at the BOL track[8] of the *MBO*, the proportion of non-western groups is "substantially larger" than "the Dutch". The report does not address the ethnicity of the students at the two higher levels HBO ('higher professional education') and WO ('academic higher education'), which suggest that the numbers of non-western students were too low to say something interesting about. The CBS numbers confirm that during the years of our study, the number of non-Western participants were indeed highest at the *MBO* education in Nijmegen. The same applies to Amsterdam. Figure 3 presents the average percentage of participation at the three levels of education by the three origins in Amsterdam and in Nijmegen.

Amsterdam has several organizations of education on MBO level, while Nijmegen has only two: an ROC (*regionaal opleidingencentrum*, 'regional training centre' for secondary vocational education) and an AOC[9]. As both cities have an ROC, we choose this organization of education to recruit our 18-20 year old participants. Two of the Amsterdam core participants (and 5 non-core participants) were recorded at ROC ASA. Nine other Amsterdam core participants were recorded at home of one of the research assistants. The Nijmegen students were recorded at the ROC of Nijmegen at three locations.

As said before, the Amsterdam Dutch participants without ethnic ties (A20C) were actually living in Aalsmeer, but studying in Amsterdam. Their level of education is higher than that of the other participants. One of them was studying at HBO level ('higher professional education'), while the other two were studying at WO level ('academic higher education').

---

[7] The students attending MBO in Nijmegen, but live outside Nijmegen, are not taken into account.

[8] One of two possible learning tracks in (secondary) vocational education. BOL students spend relatively more time in the classroom and less as apprentices working for an employer, compared to the other track.

[9] Abbreviation of *Agrarische Opleidingscentrum*: 'training centre for secondary vocational education in the sector of agriculture and the natural environment'

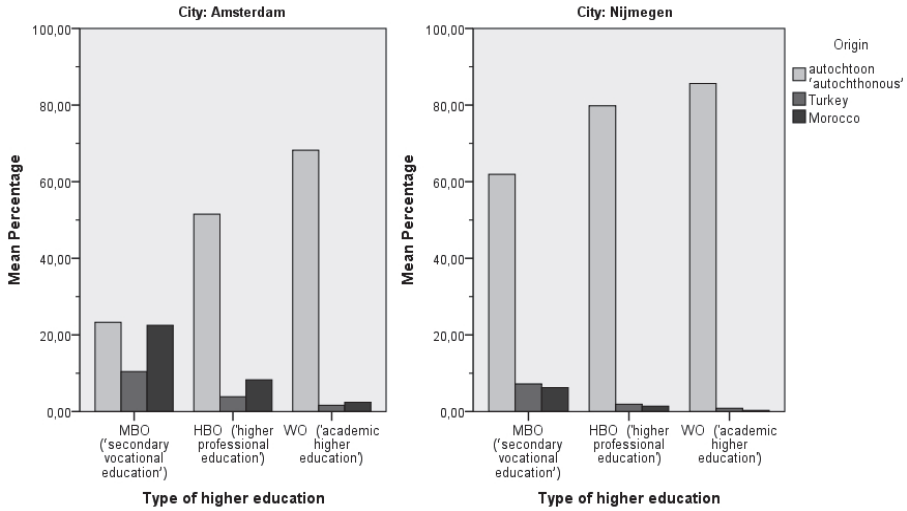


Figure 3 Populations of three types of higher education

1.6.2 Demography

1.6.2.1 Demography of areas of the schools

Amsterdam. The conversations/interviews that took place in Amsterdam were recorded in several schools in Amsterdam (three primary schools and an *ROC*), as well as at home of one research assistant. The schools are shown in Table 3 of Appendix 4 with their location, i.e. the *buurtcombinatie* 'neighborhood combinations' and the *stadsdeel* 'borough'. Amsterdam had 15 *stadsdelen* 'boroughs'[10], labeled between A and U in Map 3 in Appendix 3. Each *stadsdeel* consists of several *buurtcombinaties* 'neighborhood combinations'. In total, Amsterdam has 97 *buurtcombinaties* (code numbers from 00 to 98). Table 3 of Appendix 4 presents the number and percentage of habitants with a Turkish or Moroccan origin for each school per *buurtcombinatie* and *stadsdeel*. Just as we saw in the overall population of Amsterdam (see Figure 2), the areas of the schools also consists of more Moroccan than Turkish inhabitants.

Nijmegen. The Nijmegen conversations/interviews were recorded in several schools in Nijmegen: four primary schools and three departments (at different locations) of an *ROC* ('secondary vocational education'). Table 4 of Appendix 4 presents the schools with their *wijk* 'neighborhood' and *stadsdeel* 'borough'. Nijmegen has 9 *stadsdelen* 'boroughs', colored areas in Map 5 in Appendix 3. Each *stadsdeel* consists of several

[10] at the time of recording (2005-2007). On May 1st 2009, Amsterdam has been reorganized in 8 *stadsdelen*.

wijken 'neighborhoods'. Map 5 in Appendix 3 shows the total of 44 *wijken* (code numbers from 00 to 70). Table 4 of Appendix 4 presents the number and percentage of habitants with a Turkish or Moroccan origin in the area of each school. Just as we saw in the overall population of Nijmegen (see Figure 2), all but one areas of the schools also consists of more Turkish than Moroccan inhabitants. Only the school *RKBS Montessori* is situated in a neighborhood (Galgenveld) with slightly more Moroccans than Turks.

1.6.2.2 Demography of reported living areas of the participants

Amsterdam. With exception of the 12 year olds without ethnic ties (i.e. A12C), most participants from Amsterdam mention some information on their present and sometimes also former living area. The companion *stadsdelen* and *wijken* are presented in Table 5 of Appendix 4. For each *stadsdeel* and *wijk* belonging to the living area, the number and percentage of habitants with a Turkish or Moroccan origin is given for the year 2006.

Nijmegen. For all Nijmegen participants, some information about their living history (e.g. streets and areas) is available. The *wijken* and *stadsdelen* (belonging to the mentioned streets/areas) and the background of the reporting participants are presented in Table 6 of Appendix 4. For each *stadsdeel* and *wijk*, the number and percentage of habitants with a Turkish or Moroccan origin are given for 2006.

One 'white' Dutch boy with ethnic ties reported to live in *Beuningen* (while going to school in Nijmegen). A 'white' Dutch 20-year-old without any Turkish or Moroccan contacts, reported to live in *Mook* at the moment of the recording, but lived earlier for 8 years in Nijmegen (Brakkenstein) and also some time in *Malden*. *Mook* is part of the municipality *Mook en Middelaar* and *Malden* is part of municipality *Heumen*. *Beuningen* and *Malden* (*Heumen)* belong to *Rijk van Nijmegen*, lit. 'Land of Nijmegen'. According to the tourist information center (VVV Arnhem Nijmegen 2014), *Mook* also belongs to *Rijk van Nijmegen*. Map 6 in Appendix 3 shows *Rijk van Nijmegen* with the four mentioned municipalties/villages (*Beuningen*, *Mook*, *Malden*, and *Nijmegen*). Table 7 of Appendix 4 presents the number and percentage of habitants with a Turkish or Moroccan origin for the 3 municipalities (*Beuningen*, *Heumen* and *Mook en Middelaar*).

1.7 The main research questions of this thesis

For the study in this thesis, we will focus on seven research questions. Four of them address the roots, i.e. the origin, of certain variation patterns. Another research

question focuses on style shifting and the last two deal with the social diffusion of patterns of variation.

Roots / Origins:

1. Substrate effects

To what extent can variation patterns be related to interference from the heritage languages of speakers with Moroccan and Turkish backgrounds? Cf. (c) in Section 1.4.1 above.

2. Regional effects

To what extent can variation patterns be related to the impact of the local nonstandard varieties, i.e., to the dialect differences between the cities of Nijmegen and Amsterdam? Cf. (b) in Section 1.4.1 above.

3. Structural effects

To what extent does endogenous Dutch linguistic conditioning apply in the patterns of variation in speakers with Moroccan and Turkish backgrounds? Are conditions reset?

4. Second language acquisition effects

To what extent can variation patterns be related to properties resulting from processes of second language acquisition? Cf. (d) in Section 1.4.1 above.

Position:

5. Style-shifting effects

To what extent do speakers shift between patterns of variation, depending on the background of the interlocutor? Cf. (f) in Section 1.4.1 above.

Social spread:

6. Inter-ethnic convergence effects

Do speakers with Moroccan and Turkish backgrounds differ in their variation patterns or do they share a more global 'non-native' identity? Cf. (e) in Section 1.4.1 above.

7. Native convergence effects

Is there any evidence of spread of ethnic patterns of variation to endogenous Dutch peers? Compare Rampton's (1995) concept of crossing. Cf. (g) in Section 1.4.1 above.

1.8 Overview of the chapters

In the next chapters, we will look at variation in the realization of several Dutch by the speakers in the core corpus (cf. Section 1.6). In Chapter 2, we will examine the phoneme /z/. One of the most remarkable features of Turkish-Dutch and Moroccan-Dutch of the core corpus speakers in both cities is a characteristic dental, voiced realization of /z/. This dental realization has its origin in the various languages spoken in Morocco, and is not part of the dialectological and/or sociolinguistic patterns of variation of traditional endogenous Dutch. While six of the research questions from Section 1.7 will be addressed, the important question will be the question relating to the interference of the heritage languages of the speakers with Turkish and Moroccan backgrounds. This study thus zooms in on a phenomenon that is most likely has roots in the heritage languages of the pupils with a migrant background.

Chapter 3 will examine a phoneme that does not exist as such in the heritage languages of Turkish and Moroccan speakers, and that is at the same time subject to pronounced regional and social variation in the Dutch speech community at large. It very noticeably differentiates both urban dialects, namely the realization of the Dutch diphthong front unrounded diphthong /ɛi/. The main question is whether /ɛi/ is involved in processes of social (geographical, ethnic) redistribution. This part of the project thus focuses on a variable phenomenon that is rooted in the surrounding endogenous dialects.

Chapter 4 is focuses on the realization of a phonological contrast that is essential in Dutch (and several other modern Germanic languages and probably many other stress-timed languages), namely the long/short or tense/lax contrast in the vowel system. In the data from the core corpus, it will be studied with respect to the vowels /ɑ/ and /a:/. We will establish how speakers of syllable-timed languages such as Turkish, Arabic and Berber, who grow up bilingually with an additional stress-timed language cope with this deep typological distinction that their parents' L2 Dutch probably does not support. This study thus zooms in on a phenomenon that is rooted in second language acquisition.

As the linguistic variables in previous chapters tend to have different origins, the question arises if the features that constitute an ethnolect are coherent. Chapter 5 enters the question in which ways linguistic variables can be linked to each other and whether the resulting coherence can be explained by external (social) or internal (linguistic) factors (or both). Therefore co-variation between several linguistic variables was examined, i.e. non-accidental correlations between two or more linguistic variables in the language use of members of a specific (geographical, social or cultural) group or a speaker.

Chapter 6 recapitulates the main findings of this series of studies, draws some preliminary conclusions with regard to the main research questions, discusses the overall results, and looks ahead by identifying a number of issues for further research.

2 Ethnolectal variation in the realization of /z/ by Dutch youngsters

Edited from: van Meel, Linda, Frans Hinskens & Roeland van Hout. 2013. Ethnolectal variation in the realization of /z/ by Dutch youngsters. *Zeitschrift für Dialektologie und Linguistik* 80(3). 297–325.

Zusammenfassung

In diesem Beitrag wird eine quantitative Studie zur ethnolektalen Variation vorgestellt. Verschiedene historische sowie moderne niederländische Ethnolekte werden kurz charakterisiert. Anschließend wird eine Skizze des Korpus spontaner Sprache dargestellt, das für die Erforschung von sich entwickelnder türkischer und marokkanischer Ethnolekte des modernen Niederländischen zusammengestellt wurde. Die linguistische Variable, die im Mittelpunkt steht, ist die Realisierung von /z/ in Wortanfangsposition, mit besonderer Berücksichtigung der Prozesse der Dentalisierung und Verhärtung (oder Verstimmlosung). Wie sich herausstellt, gibt es klare Unterschiede zwischen Amsterdam und Nijmegen, den beiden Städten in denen Daten gesammelt worden sind, im Zusammenhang mit einer allgemeinen Konvergenz zwischen den Sprechern mit marokkanischem beziehungsweise türkischem Hintergrund. Die Dentalisierung von /z/ scheint ein Substrateffekt zu sein, während allmähliche Änderungen der linguistischen Konditionierung der Verhärtung auf eine strukturelle Reinterpretation seitens der Sprecher mit marokkanischem beziehungsweise türkischem Hintergrund hinweisen. Die vorgefundene stilistische Verschiebung in der Dentalisierung, die mit dem Hintergrund des adressierten Gesprächspartners zusammenhängt, scheint das Resultat eines komplexen Aushandlungsprozesses zwischen lokalen *in-* und *out-group*-Merkmalen zu sein. Unsere vorläufige Schlussfolgerung ist, dass Verstimmhaftung als ethnolektaler sowie regionaler *Indicator* funktioniert, während sich Dentalisierung als ethnolektaler *Marker* verhält.

2.1 Introduction

Ethnolects are a relatively new area in the study of language contact and bilingualism, where they have largely been studied from an ethnographic perspective. This article presents an investigation which is rooted in a language-centered, sociolinguistic approach using quantitative tools to analyze patterns of linguistic variation.

The spontaneous speech corpus we analyzed was collected within the framework of a research project on ethnolects of modern Dutch. This project, “The Roots of Ethnolects, an Experimental Comparative Study”, concentrates on the possible emergence of two young ethnolects of Dutch, as spoken by “second-generation” migrants (by which we mean speakers born and bred in the Netherlands) of Turkish and Moroccan descent, and focuses on the cities of Amsterdam and

Nijmegen. We investigated two patterns of linguistic variation in the voiced fricative /z/ in word-initial position. First of all, we examined the occurrence of a strongly voiced, dentalized /z/, a realization that has been mentioned as a typical ethnolectal feature (see El-Aissati et al. 2005). This dentalized /z/ has noticeably more frication and therefore sounds "sharper" than the traditional Dutch variant, which is alveolar. Secondly, in standard Dutch, /z/ is involved in processes of devoicing (van de Velde, Gerritsen & van Hout 1996). The question is how speakers with Moroccan and Turkish backgrounds embed those processes in their own speech, in particular as the voicing of /z/ is determined by its phonological context. In this article we try to find out whether ethnolectal varieties of Dutch are marked by stronger devoicing and/or by a reshaping of the phonological conditioning of the voiced /z/.

The corpus we investigate contains spontaneous speech data obtained from two age groups (10–12 years, 18–20 years), from two different cities (Amsterdam and Nijmegen), and from three linguistic/cultural backgrounds (Turkish, Moroccan Arabic/Berber and Dutch). All participants except for a subset of the participants in the endogenous Dutch control group took part in three different interviews, interacting with age-matched participants of (1) Turkish, (2) Moroccan, and (3) Dutch descent. This design gives us the opportunity to answer research questions related to differences between groups (the social embedding) and linguistic conditions (the linguistic embedding), and to the way variation plays an active role in face-to-face interaction (the interactional embedding). This leads to the following six research questions in relation to the patterns of variation in word-initial /z/:

Social embedding:

1. Regional effects
   To what extent can variation patterns of devoicing be related to the impact of the local nonstandard varieties, i.e. to the dialect differences between the cities of Nijmegen and Amsterdam?
2. Inter-ethnic convergence effects
   Do speakers with Moroccan and Turkish backgrounds differ in their variation patterns or do they share a more global "non-native" identity?
3. Native convergence effects
   Is there any evidence of spread of ethnic patterns of variation to endogenous Dutch peers? Compare Rampton's (1995) concept of crossing.

Linguistic embedding:

4. Structural effects

   To what extent does endogenous Dutch linguistic conditioning (progressive assimilation in the case of devoicing, depending on the preceding context) apply in the patterns of variation in speakers with Moroccan and Turkish backgrounds? Are conditions reset?

5. Substrate effects

   To what extent can variation patterns of dentalization or devoicing be related to interference from the heritage languages of speakers with Moroccan and Turkish backgrounds?

Interactional embedding:

6. Style-shifting effects

   To what extent do speakers shift between patterns of variation, depending on the background of the interlocutor?

In this article, attention is first devoted to considerations surrounding the definition of the notion of ethnolect, including a discussion of the Dutch situation in Section 2.2. In the next section, the article zooms in on the different realizations of /z/, sketching the phonetic and phonological characteristics of the sibilants in the languages at issue. The methods are presented in Section 2.4, followed by the results (Section 2.5). The final section contains the discussion and conclusions.

2.2 Background

The definition of the notion of ethnolect and its relation to concepts such as nativized varieties (such as New Englishes across the globe) and youth language is a matter of considerable and often highly controversial debate. The term *ethnolect* was coined in the late 1970s to denote "the English of the descendants of immigrant families long after their original language is lost" (Carlock & Wölck 1981; via Wölck 2002:157). According to the glossary of Danesi (1985:118), an ethnolect is "the variety of a language that results when speakers of different ethnolinguistic backgrounds attempt to speak the dominant language (e.g. Chicano English)". Unlike dialects, koines (or "regiolects"), and other homegrown varieties of a language, and unlike most transplanted varieties,[1] ethnolects are typically not the mother tongue of the first generations of speakers. Unlike transplanted varieties and "daughter languages",

[1] Transplanted varieties are varieties of a language which are spoken in a community outside the original speech community, such as the descendants of Dutch dialects spoken in Iowa (Smits 1996) and of Hunsrück dialects of German as presently spoken in parts of Brazil.

ethnolects as products of language shift usually develop in the language area of or at least in interaction with native speakers of the dominant language. Ethnolects are not necessarily purely learners' varieties as many speakers have a command of the standard variety. In the case of these speakers, it is not a matter of not being able to, but rather of not wanting to, speak the standard variety in certain situations.

Since the essence of ethnolects is still largely unknown, most definitions of the concept of ethnolect are stipulative (rather than descriptive) in the sense that they do not describe distinctive features but are conventions about what is meant by the notion. Clyne (2000:86), for instance, has defined ethnolects as "varieties of a language that mark speakers as members of ethnic groups who originally used another language or distinctive variety". Androutsopoulos (2001:2) defined an ethnolect as "a variety of the majority language (or 'host language') which is used and regarded as a vernacular for speakers of a particular ethnic descent and is marked by certain contact phenomena". Auer (2003:256; our translation) gave the following definition: "an ethnolect is a way of speaking (a style), which is associated with one or more non-German ethnic groups by the speakers themselves or by others". Auer goes on to make a distinction between ethnolect and youth language by stating that an ethnolect (also) deals with grammar, as opposed to so-called youth language, which, despite its innovations, does not.

While Auer (2003) contrasts ethnolect with youth language in terms of different features (grammar versus innovations), Muysken (2013) describes further such oppositions. Ethnolects are said to be more or less stable and their usage at most semi-conscious, ethnicity plays an inherent role, and the features involved are phonological and syntactical. Youth languages are characterized as dynamic and conscious, ethnicity plays a dynamic role, and the features involved are usually lexical and pragmatic in nature. One can add that youth language thus seems to be more of a register or even a jargon.

Muysken notes that oppositions such as those proposed by Auer "are a bit of a caricature". Both consciousness and ethnicity are complex phenomena; we do not know how stable ethnolectal features really are. Furthermore, the range of features of both ethnolect and youth language remains to be studied. According to Muysken, it seems likely that features of these types can be situated somewhere on a scale from stable to dynamic or on a scale from ethnolectal to youth-language related.

In connection with the functional dimension, the question arises as to whether ethnolects are, in Baker's (2000) terminology, Mediums for Interethnic Communication or rather Mediums for Community Solidarity. In the latter case, ethnolects will probably function exclusively or at least mainly as in-group codes; that would mean the emblematic value of the ethnolectal variants, which are often quite

distinct from the prestigious norm, is mainly defined by their signaling of ethnic identification and solidarity. This is in line with Benor's (2010:160) concept of ethnolinguistic repertoire "as a fluid set of linguistic resources that members of an ethnic group may use variably as they index their ethnic identities".

Both originally exotic features and originally indigenous nonstandard features that spread to other ethnic minority groups have come to be referred to as multi-ethnolect features (Clyne 2000; Quist 2000; Wiese 2009, 2013). Multi-ethnolect features can also be stabilized second-language-acquisition phenomena and thus need not be specific to any ethnic group and their heritage first language.

In general, there are two distinct approaches to the study of ethnolectal variation, the language-centered approach and the ethnographic approach. Whereas the ethnographic approach conceives of language systems as infinite resources from which speakers may freely choose in order to shape or construct their identity, the language-centered approach tries to disentangle the laws, generalizations and restrictions on these resources, with a distinct emphasis on linguistic embedding.

The language-centered approach is marked by terminology such as ethnolect, *multi-ethnolect* and *multicultural variety*. Sociolinguistically, the approach, which highlights features of linguistic structure, their origin and distribution, is quantitative in the Labovian tradition. The features' patterns of use are usually viewed from a rather macro-social angle (e.g. Cheshire et al. 2011; Hoffman & Walker 2010).

The ethnographic approach, which is marked by terminology such as *style* and *(pan)ethnic style*, is typically couched within the Gumperzian understanding of style as a comprehensive and (at the same time) "weak" concept. In this conception, styles have a prototypical organization. Accordingly, "ethnic ways of speaking" are seen as "rather fuzzy phenomena with some core linguistic features and more variable features at the boundaries" (Kern 2011b:9). Styles are not determined; they are continuously being constructed as socially and interactively significant products (cf. Eckert's 2008 view that ethnolects are fluid rather than fixed entities). Yet, in contrast to the language-centered approach, language change is not a central concern in the ethnographic approach. Both reactive and initiatory uses of linguistic and paralinguistic features are analyzed. The angle is micro-social and interpretive (e.g. Benor 2010; Keim & Knöbl 2007; Kern & Selting 2011).

At least seven Dutch ethnolects can be distinguished within the period from the middle of the eighteenth century till the last quarter of the twentieth century: Jewish Dutch (which started to emerge around the middle of the eighteenth century; cf. den Besten & Hinskens 2005:289–293), Indonesian Dutch (1900–), Surinamese Dutch (1900–), Moluccan Malay Dutch (1920–), Antillean Dutch (1950–), Moroccan Dutch (1970–) and Turkish Dutch (1970–; cf. Hinskens 2011; Muysken 2013).

Surinamese Dutch and Indonesian Dutch arose outside of the Netherlandic area and are considered colonial varieties. Moluccan Malay Dutch and Antillean Dutch may have originated outside of the Netherlands as well, but the majority of immigrants came to the Netherlands primarily as speakers of a range of local languages. Jewish Dutch, Moroccan Dutch and Turkish Dutch arose in the Netherlands and are immigrant varieties. Most ethnolects are spoken in urban areas except Moluccan Malay Dutch, which is spoken mostly in smaller provincial towns (Muysken 2013).

2.3 The realization of /z/

Two earlier studies report possible ethnolectal variants of /z/. In their impressionistic sketches of Moroccan Dutch by female speakers of the first and second immigrant generation, El Aissati and his coauthors (2005:155, 162) mention strongly voiced and "scherp" (lit. 'sharp') realizations of [z]. Hinskens (2011) contains a preliminary analysis of length and voicing variation in the realization of /z/ in the same ethnolect corpus we investigate here. In this section we overview /z/ in the languages involved.

2.3.1 Dutch

Dutch has categorical final devoicing of obstruents in syllable final position. The realization of /z/ is variable in the onset position. In this section, we first describe /z/ and its voiceless phonemic counterpart /s/ in Standard Dutch and then examine how these sounds are realized in the dialects of Amsterdam and Nijmegen, as the speakers in our sample live in these cities (see Section 2.4.1.).

2.3.1.1 Standard Dutch

Standard Dutch has voiced and voiceless alveolar fricatives. These fricatives can be considered as phonemes as there are a few minimal pairs such as in (1).

(1) *zus* 'sister' – *sus* 'soothe-1sg'
zee 'sea' – *C* 'cee'

The voiced phoneme /z/ has voiceless /s/ as a counterpart. In practice, speakers of Netherlandic Dutch make this distinction less and less often and use the voiceless [s] instead of the voiced [z]. Van de Velde, Gerritsen & van Hout (1996) show that the devoicing of /z/ had already started between 1965 and 1980 and became even stronger thereafter. Especially north of "the big rivers" Maas/Meuse and Rhine, there is a tendency to devoice fricatives (Cohen et al. 1971). Since there are only a few minimal pairs, in most cases the pronunciation has no consequences for interpretation of the word.

Due to progressive assimilation, voiced fricatives devoice after a voiceless obstruent (Booij 1995:58). This occurs both within words as in (2) ([2a] compound word, [2b] particle verb) and across (grammatical) word boundaries as shown in (3).

(2)	a. *zitzak*	/zɪtzɑk/	[zɪtsɑk]	'bean-bag'
	b. *opzoeken*	/ɔpzukən/	[ɔpsukən]	'look up'
(3)	*gaat zomaar*	/χɑt zomar/	[χɑt-somar]	'goes-3sg just (like that)'

2.3.1.2 The local dialect of Amsterdam

According to Daan (1948), /z/ at the beginning of a word is always realized as [s], e.g. *zuur* [su:r] 'sour'. In *plat Amsterdams* (broad Amsterdam dialect), this sound also occurs in the middle of words, e.g. *ruzie* [ru:si:] 'quarrel'. Schatz (1986:61) states that the devoicing of /v/, /z/ and /g/ is "the most pervasive feature in Amsterdam". It is part of stereotyping with respect to the Amsterdam dialect, i.e. it is one of the first phenomena imitated when mimicking someone from this city. According to Schatz, the devoicing of /z/ occurs in all sectors of society. Brouwer (1989) studied the use of the devoiced /z/ in Amsterdam. Her participants, all born and raised in Amsterdam, used many devoiced variants with an index value of 69.13 in casual speech (100 is complete devoicing). Even in more careful speech – reading word lists aloud – the index score was over 50 (more devoicing than voicing).

2.3.1.3 The local dialect of Nijmegen

Just as in the local dialect of Amsterdam, yet more variably, the dialect of Nijmegen has a strong tendency to devoice fricatives in syllable onset position (van Hout 1999). Van Hout (1989) also studied the use of the devoiced /z/ in a word list reading task in Nijmegen. His participants, all born and raised in Nijmegen, produced variants with an index value of 46 on a scale of 0–100 (100 is complete devoicing). This devoicing of fricatives places the Nijmegen dialect in a remarkable position, as it is located within an area where the voicing opposition still stands in the dialects.

2.3.2 Turkish

Several authors give different specifications of the position of the tongue in pronouncing Turkish /s/ or /z/. Swift (1963:11) described Turkish /s/ and /z/ as "alveolar grooved fricatives". This is followed by a more detailed description: "/s/ is a voiceless fricative made with the tongue grooved against the alveolar ridge behind the upper teeth, much like, but slightly more forward than the English 's' in *son* or *pass*" and "/z/ is the voiced 'equivalent' of /s/, much like the English 's' of *phase*".

According to Göksel & Kerslake (2005), these fricatives are denti-alveolar, according to Kornfilt (1997) they are dental or alveolar, and both Comrie (1997:885)

and van der Hulst & van de Weijer (1991:13) state that the Turkish fricatives /s/ and /z/ are dental.

Several studies (Comrie 1997:891; Kornfilt 1997:492; Lewis 2000[1967]:2) mentioned that native words do not have [z] in initial position, except from some onomatopoeia or words from children's language, like *zır zır* 'in an incessant, nerve-racking way', *zıplamak* 'to bounce/jump (up and down)' and *zıpzıp* '(a) marble (a child's toy)'. However, Turkish has many loanwords, for instance from Arabic, beginning with /z/.

2.3.3 Moroccan languages

Morocco is commonly characterized as a multilingual country. Most Moroccans speak Moroccan Arabic or Berber languages as their mother tongue(s) (El-Aissati et al. 2005:150; El-Aissati & E-rramdani 2001:63).

Moroccan Arabic consists of several dialects that are mutually intelligible. The Berber languages can be divided into three main groups: 1. Tarifit or Riffian or Rif Berber, 2. Tachelhit or Sous(s) Berber, 3. Tamazight or Central/Middle Atlas Berber.[2] These three groups are not mutually intelligible.

El Aissati and coauthors (2005:150) state that at a rough estimate sixty percent of "the Moroccans in the Netherlands" speak Tarifit, ten percent speak Tachelhit and the remaining thirty percent speak a Moroccan Arabic dialect as their mother tongues.

In the next two sections, /z/ in Moroccan Arabic and in the Berber languages Tarifit and Tachelhit is discussed.

2.3.3.1 Moroccan Arabic

The Moroccan Arabic /s/ and /z/ are apical or dental fricatives (Abdel-Massih 1973; Harrell 1962, 1965). According to Harrell (1962:4), they

> differ from the corresponding English sounds in that the tip of the tongue touches or approaches the back surface of the upper front teeth. Corresponding sounds in English are made with the tip of the tongue touching or approaching the gum ridge above the upper front teeth.

In contrast to Dutch and Turkish, Moroccan Arabic has both a plain and an emphatic version of /s/ and /z/. These emphatics, /sˤ/ and /zˤ/, have phonemic status as (near) minimal pairs like that shown for /s/ and /sˤ/ in (4) exist.

---

[2] We use the names and spellings listed in Ethnologue (URL: <http://www.ethnologue.com/show_country.asp?name=MA>; accessed 21 October 2010): Tarifit, Tachelhit and Tamazight.

(4) *sif* 'sword' – *sˤif* 'summer'

Emphatic consonants are also referred to as pharyngealized consonants (see for example Thelwall & Sa'Adeddin 1990) and sometimes "velarized" or "flat" (Abdel-Massih 1973:5). When pronouncing an emphatic consonant, there is greater muscular tension in the mouth and throat, and the tongue is raised toward the roof of the mouth (Harrell 1965:5). This causes an articulation that is not just dental, but also velarized or alveolarized (Abdel-Massih 1973:5).

2.3.3.2 Berber

The dialect variation in Berber phonological systems is large (Kossmann & Stroomer 1997:463). We will discuss the two Berber dialect groups that are spoken by Moroccan Dutch: Tarifit and Tachelhit.

2.3.3.2.1Tarifit

Just like Moroccan Arabic, Tarifit has emphatic (pharyngeal) /sˤ/ and /zˤ/ in addition to the plain /s/ and /z/. The emphatic consonants are phonemic as McClelland III (2008) shows with several (near) minimal pairs. Example (5) presents such a minimal pair. Note that the emphatic phoneme /zˤ/ alters /æ/ so that it is pronounced as [a].

(5) *zwæ-nt* 'they crossed' – *zˤwa-n-t* 'they hurt him' (McClelland III 2008:63)

An obstruent cluster as a whole is typically either voiced or unvoiced, which implies that Berber has assimilation rules, cf. McClelland III (2008:58).

2.3.3.2.2Tachelhit

Kossmann & Stroomer (1997) state that Tachelhit Berber has plain and pharyngeal (i.e. emphatic) dental fricatives. Just as in Moroccan Arabic and Tarifit, the emphatic consonants are considered phonemic.

2.4 Methods

The empirical study presented here is based on quantitative analyses of data collected from a stratified sample of speakers. The speaker sample as well as the different types of output collected for every single speaker fit a factorial design, i.e. one in which several (extra-linguistic and linguistic) factors have been simultaneously manipulated or controlled systematically. In this respect, our work is comparable to ethnolect studies such as those by Hoffman & Walker (2010) and Cheshire et al. (2011).

2.4.1 Speakers

Data from 51 participants were selected for analysis. The groups were controlled for background (Moroccan, Turkish and Dutch), age (10–12 versus 18–20 years old), residence (Amsterdam versus Nijmegen) and, in the case of the speakers with a Dutch background, the presence or absence of regular contact with speakers of Turkish and Moroccan descent. Table 1 shows the speaker design.

Table 1 Overview of the number of participants

Background	Moroccan		Turkish		Dutch			
Inter-ethnic ties?	Yes		Yes		Yes		No	
Age (years)	10-12	18-20	10-12	18-20	10-12	18-20	10-12	18-20
Amsterdam	3	4	3	4	2	3	3	3
Nijmegen	3	3	3	3	3	3	4	4

The participants with a Moroccan or Turkish background were born in the Netherlands and have at least one parent who immigrated from Morocco or Turkey respectively. All participants grew up in their place of residence (either Amsterdam or Nijmegen). For the sake of brevity, we refer to the participants with Moroccan and Turkish backgrounds as Turkish Dutch and Moroccan Dutch. The two groups with an entirely Dutch background will be referred to simply as Dutch or native Dutch (whether with or without interethnic ties).

Table 2 Self-reported language skills; for all subsets of speakers, the languages are ordered identically, starting with Dutch, i.e. they do not necessarily reflect the order given by the speaker(s); (English, German, French and Spanish are foreign languages learned in secondary education)

Background	Age (years)	Speakers (*n*)	Language skills
Turkish-Dutch	10-12	6	Dutch, Turkish
	18-20	2	Dutch, Turkish
		4	Dutch, Turkish, English
		1	Dutch, Turkish, English, German, Spanish
Moroccan-Dutch	10-12	2	Dutch, ‘Moroccan’
		1	Dutch, Moroccan Arabic
		2	Dutch, Berber
		1	Dutch, Berber, “a little Moroccan Arabic”
	18-20	1	Dutch, ‘Arabic’
		1	Dutch, Moroccan Arabic, English
		1	Dutch, Moroccan Arabic, English, German
		1	Dutch, Moroccan Arabic, English, French, Spanish
		2	Dutch, Berber, English
		1	Dutch, Berber

Unanticipated complications during the fieldwork sessions mean only scant information is available about the language skills and linguistic profiles of the speakers in our sample. The “linguistic profile” notion refers to such issues as the

domains of language use, network-specific linguistic practices and the like. For 14 of the 51 speakers in the sample studied, we have no information whatsoever about their linguistic profiles; seven of them were endogenous Dutch, four Turkish Dutch and three Moroccan Dutch. All Turkish Dutch speakers reported knowledge of Turkish and all Moroccan Dutch speakers reported knowing one or more of the Moroccan languages, as shown in Table 2.

Several things need to be kept in mind when reading Table 2. First, "language skills" was usually treated as a *yes/no*-variable by the informants and there is no reliable information regarding the speakers' relative proficiencies in the languages mentioned or in the varieties of the languages at issue. Second, certain languages were sometimes not precisely identified (e.g. "Moroccan" sometimes refers to Moroccan Arabic and sometimes to Berber, and in yet other cases probably both), at times because control of the language in question might have been considered politically loaded – as can be the case for Kurdish. The fact that precise information was not always asked for by the fieldworkers further complicates the picture. For all these reasons, we refrain from using this information in our analyses.

2.4.2 Material and data

The participants had relatively free conversations of about one hour each with one partner per conversation. The conversation partner was either a peer from their own background (an in-group conversation) or a peer with another background (an out-group conversation). The speakers selected for the present study talked in at least one in-group situation and two different out-group situations (i.e. they had conversations with speakers from all three backgrounds). This doesn't apply to those endogenous Dutch participants who had no (or at best very weak) interethnic ties and only conversed with other endogenous Dutch peers who themselves had no interethnic ties. They are the control group for the endogenous Dutch participants who did have such ties, to which they can be compared in the in-group condition. The majority of the speakers attended the same school as their conversation partners, and many of them were classmates (especially so in the case of the 10–12 year olds). Preferably, the interviewer was only present at the beginning and the end of the exchange to ensure a more natural conversation. However, especially so in the case of the 10–12 year olds, additional guidance was needed to keep them talking for one hour (e.g. suggesting topics to talk about, introducing card games).

The conversations were recorded on a Marantz Professional CD recorder CDR300. Ten to fifteen minutes of each recording were transcribed using the multimedia annotator Elan (cf. Brugman & Russel 2004). A second person checked the transcriptions.

From each conversation, 30 /z/-words were selected per speaker, starting with the checked part of the transcriptions. If there were not sufficient /z/-words in the checked part, first the unchecked transcriptions were scanned and ultimately non-transcribed parts were listened to, as needed.

The first three minutes of each conversation were skipped, as we expected the participants to need some time to forget they were being taped and to stop paying attention to their way of speaking. We carefully selected only those realizations that gave unambiguous information on the linguistic variable based on the following criteria:

1. Only prosodic words that begin with a /z/ were selected, hence including compounds which have a /z/ at the beginning of the second element and particle verbs. See (6a–c) for examples; (6d) does not match criterion 1.
2. /z/ must be followed by a full vowel, hence not by a schwa. The examples in (7) were not selected.
3. No /z/ following a sibilant was selected.
4. To ensure the data were not biased by a specific high-frequency word, a word was selected at most three times for a given speaker in a given conversation. A homophonous word from a different word class was counted as a different word.[3] For example *zeg* can be selected six times per speaker per conversation: three times as a verb (say-1sg/2sg/imp) and three times as an interjection (say-int). Another example is *zat* which occurs in three word classes: 'sit-past' (verb), 'drunk' (adjective) and 'plenty of' ([indefinite] numeral), and might therefore be selected nine times.
5. No /z/ following a pause was selected.
6. Words with an orthographical <z> which are pronounced with [s] according to the standard norms were excluded (e.g. *zestig* /s'ɛstəχ/ 'sixty', *zeventig* /s'evə(n)təχ/ 'seventy').
7. Words that were read, for example from newspapers and magazines, were not taken into account.
8. Words that were uttered in an (intentionally) conspicuous way were not selected either. For the most part these were imitations.

(6) a. *zaterdag* 'Saturday'
b. *koopzondag* '≈ Sunday opening hours'
c. *opzoeken* 'look up'
d. *gezellig*' ≈ pleasant' > /z/ is not initial

[3] Seven word classes were distinguished: noun, verb, pronoun, adjective, adverb, interjection and "remaining word classes". The remaining word classes include cardinals (e.g. *zes* 'six'), ordinals (e.g. *zesde* 'sixth'), quantifiers (e.g. *zoveel* 'that much'), conjunctions (e.g. *zodat* 'so [that]') and prepositions (e.g. *zonder* 'without').

(7) a. *zwaar* 'heavy' > /z/ is followed by a sonorant
b. *z'n* [zən] 'his' > /z/ is followed by a schwa

All selected words were coded for the preceding context: (1) obstruent, (2) sonorant, and (3) vowel. The preceding items were classified in terms of their actual pronunciation. For example the word wat [ʋɑt] belongs to context (1) as /t/ is an obstruent, but if the /t/ were not pronounced ([ʋɑ]), the word would belong to context (3). For each context at least five /z/-words were selected. If the thirty words selected earlier did not contain enough /z/-words for a given context, the recording was searched for more words for that category until five instances were found. For some speakers, this results in more than thirty selected words per conversation. On the other hand, for some participants less than thirty words (minimum: 18) were analyzed because they did not use enough /z/-words meeting the criteria. The average number of words selected was twenty-nine per conversation per speaker.

2.4.3 Variants and coding

As mentioned in the introduction, we investigated two patterns of linguistic variation in the voiced fricative /z/ in word-initial position: voicing of /z/ and the dentalization of voiced /z/. The transcription and coding of the realization of /z/ was performed by the first author. Two variants were distinguished when rating voicing: if the /z/ sounded voiceless it was coded as 0; if the /z/ was voiced it was coded with 1. The voiced variants were coded as either non-dentalized (value 0) if they sounded like a endogenous Dutch voiced /z/ or as dentalized (value 1) if they sounded "sharper" than this. The realization of dentalized /z/ is considerably more fronted, i.e. dental, than the "regular" Dutch /z/ (cf. Harrell's 1962 description of Moroccan Arabic /z/ and English /z/ quoted in Section 2.3.3.1).

To test the value of these ratings, the degree of agreement was tested by having 181 realizations checked by a second rater, a trained phonetician. These realizations were selected randomly from the data of fourteen of the 20-year-olds. Only realizations from the transcribed part of the in-group conversations were taken. Furthermore, only words in which /z/ was preceded by a vowel or sonorant were selected. This resulted in nine to eighteen tokens per speaker. Cohen's kappa showed an acceptable agreement value of .723 for the variable property voicing. Dentalization was analyzed for the 90 voiced realizations. The agreement value for dentalization, Cohen's kappa = .707, was acceptable as well.

We also carried out measurements on the same sample with the Praat program (Boersma & Weenink 2010). The harmonics-to-noise ratio (HNR; cf. Boersma 1993) was measured for voicing. Two measurements were taken for dentalization: (1) centre of gravity (CoG), and (2) sound exposure measurement (E). The sound

exposure measurement was calculated by taking the sound exposure (in Pascal-squared seconds) between 6,000 and 10,000 Herz and dividing it by the sound exposure between 2,000 and 6,000 Herz. Both measurements capture how a sound is weighted across low to high frequencies. The higher the centre of gravity and sound exposure measurement values, the more dental a given realization of /z/ is.

Pearson correlations were carried out to test the relationship between the first author's codings and the measurements. There was a significant positive correlation between the coding on voicing and the harmonics-to-noise ratio: .537 ($p = .000$). Significant positive correlations were also found for the coding of dentalization with the centre of gravity and the sound exposure measurement: .254 ($p = .016$) and .241 ($p = .031$), respectively. However, the correlations were low, which may well be because individual differences between speakers were not taken into account. We therefore applied linear mixed-model analyses to the sample with the speaker as the random variable and the codings of the first author as the fixed variable. The first analysis showed a significant difference for the harmonic-to-noise ratio between devoiced and voiced /z/ tokens – $F(1, 163.288) = 54.704$, $p = .000$. In other words, the instances of /z/ that were coded as voiced had significantly higher HNR scores than those coded as voiceless. For the centre of gravity measurements, only the four speakers from the sample who used both voiced dentalized and voiced non-dentalized variants were taken into account. This analysis also returns a significant difference – $F(1, 28.889) = 5.916$, $p = .021$. This means that the instances of /z/ that had been coded as dentalized indeed showed a systematically higher centre of gravity than those coded as non-dentalized. Sound exposure measurement could only be calculated for three of the speakers who had used dentalized /z/. The analysis with this relative energy measurement gave an even more strongly significant effect than did the centre of gravity scores – $F(1, 23.133) = 12.698$, $p = .002$. We consider these findings as calibrations which show that the perception and coding of the dependent variables are acoustically adequate.

2.4.4 Data analyses

First, the two endogenous Dutch groups were compared on the basis of data from the in-group conversations (since those Dutch without interethnic ties only participated in-group conversations). Next, both the in- and out-group conversations of those Dutch with interethnic ties, the Turkish Dutch, and the Moroccan Dutch were analyzed. ANOVAs (GLM, repeated measures) with one linguistic factor, viz. preceding context, and four external factors (one geographical, two social and one regarding style shifting) were performed. The results are presented in Table 3.

Table 3 Factors

External factors			Linguistic factor
Geographical	Social	Style shifting	
City: Amsterdam vs. Nijmegen	Background of speaker Age: 12 vs. 20	Background of interlocutor	Preceding context
	Inter-ethnic ties Dutch group		

The “background of the interlocutor” as well as the linguistic factor “preceding context” were treated as within-subject factors. As mentioned, only voiced realizations of /z/ were used to code dentalization. As two 20-year-olds from Amsterdam, one from the Turkish Dutch and one from the Moroccan Dutch group, did not realize any voiced /z/ in one of the three contact situations (namely “background of the interlocutor”), they were excluded from the analysis.

For additional analyses on the level of the individual speakers, where we looked at frequencies of occurrence, we used chi-squared tests (or Fisher’s exact tests when the expected values were too low).

2.5 Results

We first discuss the findings for dentalization, before zooming in on the considerably more complex results for voicing. Given the research design, we needed to test effects for (1) background (Dutch/Turkish/Moroccan), (2) the age and (3) city of origin of the speaker, (4) the background of the interlocutor (Dutch/Turkish/Moroccan), and (5) the linguistic context (preceding obstruent/sonorant/vowel). A separate step was to test whether differences exist between the two endogenous Dutch groups (those with versus those without interethnic ties). We start the analyses for both dentalization and voicing by comparing the two endogenous Dutch groups.

2.5.1 Dentalization

The /z/ realizations by endogenous Dutch participants with interethnic ties (i.e. with Turkish Dutch and/or Moroccan Dutch friends) showed zero dentalization, as was also observed for the endogenous Dutch participants with no interethnic ties (i.e. no regular contact with Turkish Dutch and Moroccan Dutch). This was in sharp contrast to the Moroccan Dutch and Turkish Dutch participants. The proportion of dentalized tokens among the Turkish Dutch speakers was 0.1043 (10.43 percent) and that for the Moroccan Dutch was 0.1051. Given the complete absence of dentalized realizations produced by the endogenous Dutch group, we restricted further analysis to the Turkish Dutch and Moroccan Dutch groups.

The speaker's age group did not show a significant main effect, nor was age involved in any significant interaction effects. The same holds for the linguistic context. Accordingly, the discussion of variation in /z/ dentalization can be confined to the effects of the remaining three independent variables. The differences related to city of residence and the speakers' and the interlocutors' backgrounds are depicted in Figure 1.

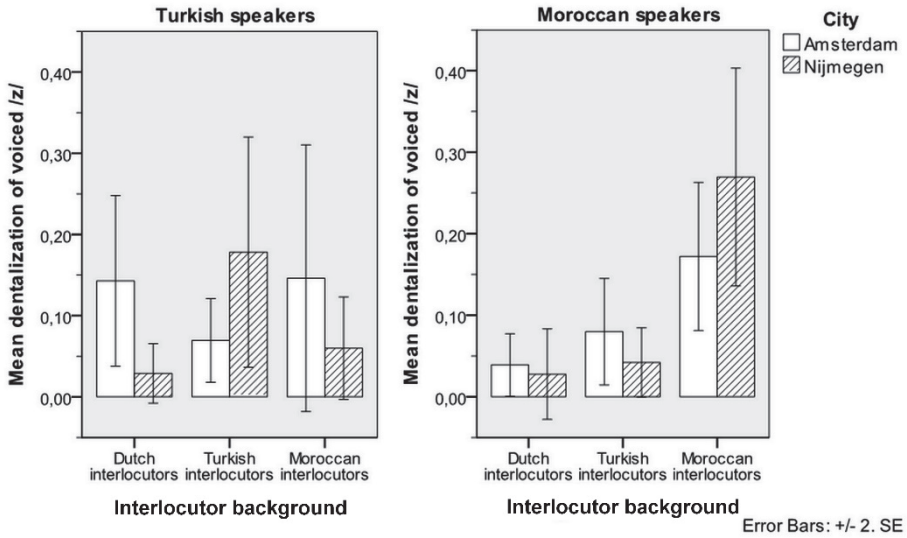


Figure 1 Mean dentalization of voiced /z/ by Turkish speakers and Moroccan speakers, broken down by city and interlocutor

The speakers' city does not return a significant main effect. There is no general difference between Nijmegen and Amsterdam. Dentalization in the realizations of /z/ appears to depend on the background of the speaker – $F(2,25) = 5.644$, $p = .011$ – and the background of the interlocutor – $F(2,50) = 6.722$, $p = .003$ – but these effects are qualified by additional interaction effects. There is a significant interaction effect between the backgrounds of the speaker and the interlocutor – $F(4,50) = 6.192$, $p = .000$ – as well as a three-way interaction between the background of the speaker, of the interlocutor and the city – $F(4,50) = 4.060$, $p = .007$.

The interaction between the backgrounds of speaker and interlocutor shows a sharp difference in style shifting between the two groups. For the Turkish Dutch speakers, their mean overall dentalization index is 0.0859 in interactions with endogenous Dutch interlocutors, 0.1239 in conversations with other Turkish Dutch and 0.1030 in interactions with Moroccan Dutch, showing no style shifting. The pattern is different for the Moroccan Dutch speakers: their mean overall dentalization index is 0.0334 in interactions with endogenous Dutch interlocutors, .0610 in

conversations with Turkish Dutch and 0.2208 in interactions with other Moroccan Dutch. We analyzed the style-shifting pattern in more detail on the individual level. Not all speakers used dentalized /z/ with all three kinds of interlocutors. Six speakers did not use a dentalized /z/ when speaking with endogenous Dutch interlocutors. Two of the Moroccan Dutch speakers did not use it with Turkish Dutch interlocutors. However, four Turkish Dutch speakers used dentalized /z/ only with Turkish Dutch interlocutors, while three Moroccan Dutch speakers used it only with Moroccan Dutch ones. One Moroccan Dutch and one Turkish Dutch speaker, both from Amsterdam, did not use a dentalized voiced /z/ at all. Although both speakers claimed to know either Moroccan Arabic or Turkish, they reported speaking only Dutch with their family and friends. It is not clear to what extent a relationship might exist between their exclusive use of alveolar /z/ and their only speaking Dutch with their friends and relatives.

Chi-squared tests (Fishe r's exact tests) showed that five speakers, four of whom were Moroccan Dutch, shifted significantly depending on the background of the interlocutor, producing more dentalized /z/ tokens in intra-ethnic group situations. Dentalization is a style-sensitive linguistic variable. Moroccan Dutch shifted more, but that does not preclude the possibility that individual Turkish Dutch switch depending on the interlocutor.

The significant three-way interaction points to a difference in the behavior of Turkish Dutch speakers from the two cities, although the differences were not large. The Amsterdam Turkish Dutch exhibited about the same level of dentalization for both endogenous Dutch and Moroccan Dutch interlocutors, while it was somewhat lower for Turkish Dutch interlocutors. The pattern for the Nijmegen Turkish Dutch was a mirror image of this, with a low dentalization index for endogenous Dutch and Moroccan Dutch interlocutors and a relatively high one for fellow Turkish Dutch. The Moroccan Dutch speakers displayed the same pattern in both cities. Moroccan Dutch in both Amsterdam and Nijmegen had the lowest dentalization indexes in their conversations with endogenous Dutch interlocutors and the highest ones in their interactions with fellow Moroccan Dutch, with those for Turkish Dutch interlocutors roughly in between. Dentalization seems to be a stylistic marker for the Moroccan Dutch speakers in particular.

2.5.2 Voicing

Is there a difference in devoicing between the two endogenous Dutch groups, the group with interethnic ties versus the group with no such ties? No effects whatsoever were found for group or city. That means that same level of devoicing was found in Amsterdam and Nijmegen, the behavior of both endogenous Dutch groups was the

same. Two main effects were significant. The strongest by far was linguistic context – $F(2,34) = 68.602$, $p = .000$ – where the obstruent context was sharply distinguished from the other two (sonorant and vowel). The significance of age – $F(1,17) = 6.595$, $p = .020$ – reflected more devoicing in the older age group, an effect discussed in more detail later in this section.

We applied an overall analysis to the endogenous Dutch, Turkish Dutch, and Moroccan Dutch participants who were recorded in interviews with three different interlocutors (the endogenous Dutch subgroup with interethnic ties). The strongest effect by far was linguistic context – $F(2,50) = 126.678$, $p = .000$; partial eta squared = 0.835. The second strongest effect was the main effect for group, which was the background of the participant (partial eta squared = 0.436). The impact of both main effects is illustrated in Figure 2.

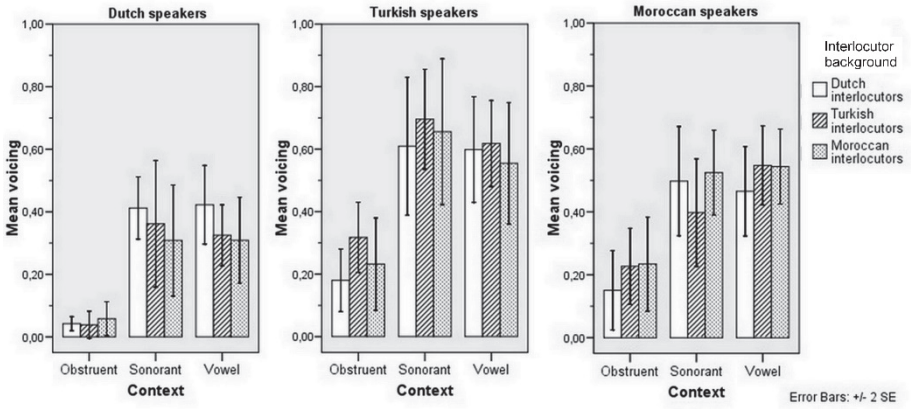


Figure 2 Mean voicing of /z/ for Dutch, Turkish and Moroccan speakers, broken down by context and interlocutor

The bars in Figure 2 show the distinction between the post-obstruent context (strong devoicing) and the other two (more voicing). At the same time, the Moroccan Dutch and Turkish Dutch participants displayed more voicing than the endogenous Dutch participants, indicating that they are different and perhaps lagging behind in the ongoing process of devoicing in Dutch. The Turkish Dutch participants obtained the highest voicing scores, a result that perhaps requires an explanation grounded in the properties of the participants' first language, or in the situational setting, or both. This issue obviously requires further research. In fact, the most striking result is the distinction between the linguistic contexts, as the post-obstruent context is actually voiceless for the endogenous Dutch group.

The /z/ realizations by the endogenous Dutch speakers showed hardly any voicing (0.0453) following an obstruent, which does not come as a surprise given the fact that progressive voice assimilation – discussed in Section 2.3.1. – is a categorical rule in all varieties of Dutch. The Moroccan Dutch and Turkish Dutch speakers had voicing indexes of 0.2037 and 0.2435, respectively; in other words, in about one fifth and one quarter of the cases (respectively) they voiced /z/ and hence – regressively – the cluster,[4] which is rather exotic to the Dutch ear.

Chi-squared tests (Fisher's exact tests) revealed that six speakers (two of whom were 20 years old, three of whom live in Nijmegen, one of whom has a Turkish and two a Moroccan background) did not make a distinction between the three preceding environments. The three endogenous Dutch speakers produced mainly voiceless tokens in all three environments, as did the Turkish Dutch speaker. The twelve-year-old Moroccan Dutch speaker seemed to have no preference for either voiced or voiceless instances, while the 20-year-old Moroccan Dutch produced mainly voiced instances. A blurring of the linguistic condition of "preceding obstruent" with the other two environments with regard to the use of voiced variants seemed to be related to having a Moroccan background. The blurring of the voice contrast itself seemed to be more of an indicator of a Dutch background.

Figure 2 also makes clear that the degree of voicing in the postvocalic and post-sonorant contexts showed highly similar patterns for the significant main effects and the interaction effects they are subject to, in contrast to the post-obstruent context. The only common effect – and in fact the only significant effect for the voicing of /z/ following obstruents – was that for the speaker's background. We restricted further analyses to the sonorant and vowel contexts.

The effects of the background of the speaker and the interlocutor in relation to city are depicted in Figure 3. There were significant main effects for the background of the speaker – $F(2,25) = 9.070$, $p = .001$ – and the city – $F(1,25) = 6.783$, $p = .015$. The mean voicing index for the endogenous Dutch speakers' /z/ realizations in these contexts was .3636, for the Moroccan Dutch it was 0.4961 and for the Turkish Dutch it was 0.6220, a ranking we discussed where the obstruent context was included. Figure 3 makes clear that there is no main effect for the background of the interlocutor. There was a significant interaction between city of residence and the preceding context as well – $F(1,25) = 6.537$, $p = .017$. While the /z/ voicing indexes for the Amsterdam speakers were 0.4037 following sonorants and 0.4468 following vowels, the reverse order applied to the Nijmegen speakers, whose voicing indexes were 0.6138 and 0.5477, respectively. So in Amsterdam, a preceding vowel triggers

[4] Alternatively, both groups use intermediate variants in almost half of the cases – or a combination of these extremes.

more voicing than a preceding sonorant, whereas in Nijmegen the order of the effects is reversed. The differences are not large, but perhaps interesting enough to pursue in further research.

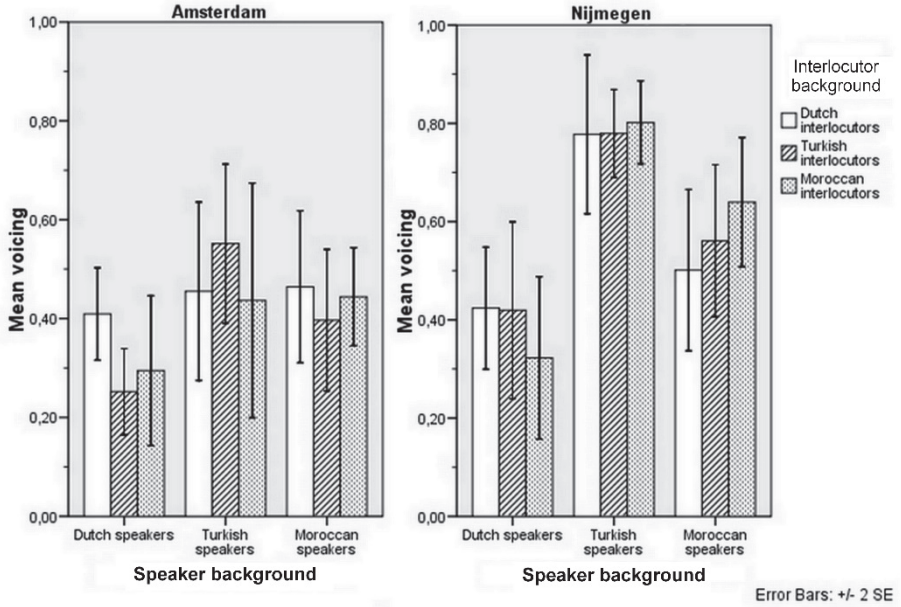


Figure 3 Mean voicing of /z/ for Dutch, Turkish and Moroccan speakers, broken down by backgrounds of speaker and interlocutor and by city; context is restricted to preceding vowel and sonorant

Figure 4 shows the histograms for age in relation to city and background of the speaker. Significant interactions were found between city and age – $F(1,25) = 15.215$, $p = .001$ – and between the speaker's background and age – $F(2,25) = 9.404$, $p = .001$. For the twelve-year-olds, the voicing indexes for the Amsterdam and the Nijmegen speakers were not too distinct (0.5873 and 0.4839, respectively); for the twenty-year-olds, the ordering was not only reversed, but the difference was large, with the Amsterdam speakers having an average voicing index of 0.3075 and the Nijmegen speakers 0.6775, as can be seen in the figure. The /z/ voicing index for the endogenous Dutch twelve-year-olds was 0.2467, that for their Moroccan Dutch peers was almost twice as high (0.4644), while the index for the Turkish Dutch twelve-year-olds was almost twice as high again (0.8390). The pattern for the twenty-year-olds was quite different (endogenous Dutch = 0.4610, Moroccan Dutch = 0.5232, Turkish Dutch = 0.4360).

Two complex interaction effects remain. The first one was the interaction between the background of the speaker, of the interlocutor, and age – $F(4,25) =$

2.669, p = .043. The second was the interaction between background of the speaker, of the interlocutor, age and city – F(4,25) = 2.801, p = .036. These effects are modest in size, and they make clear in particular that age is an effect that needs to be studied in greater detail to understand which processes of change are going on and to determine which age groups are leading these changes and how age cohorts differ.

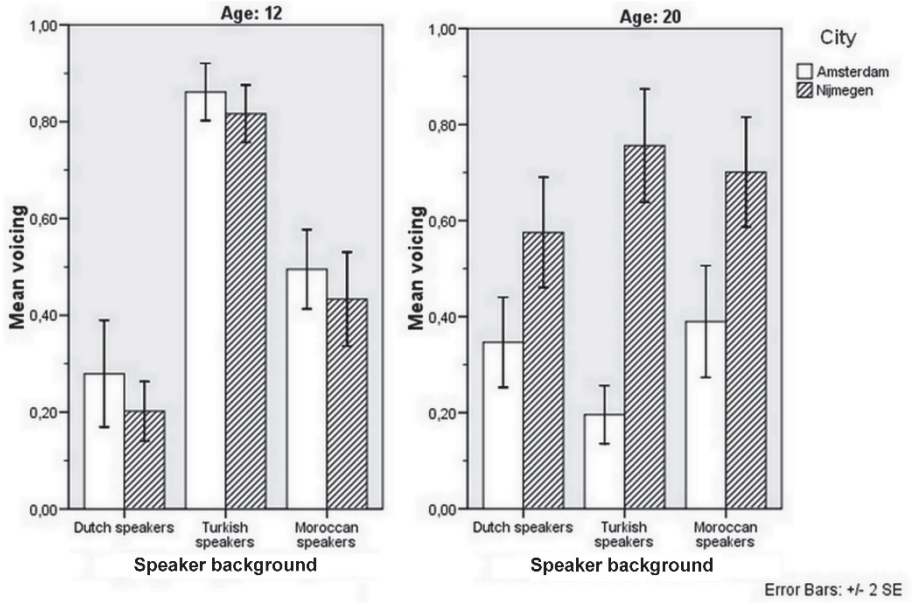


Figure 4 Mean voicing indexes for 12 and 20-year-olds, broken down by city and background of the speakers; context is restricted to preceding vowel and sonorant

2.6 Discussion and conclusion

Table 4 summarizes the results we found for the two variables (dentalization and voicing) for each of the three types of embedding we distinguished. We then took the concepts of marker and indicator, in the sense of Labov (1971, 1972a), who distinguished between them on stylistic grounds, and applied them to the variables in Table 4. Markers show stylistic shift, indicating that the linguistic variable is manipulated by speakers. We saw that the Turkish Dutch and Moroccan Dutch participants adapted their speech to the background of the interlocutor, an interactional stylistic shift that shows that ethnolect speakers adapt their linguistic behavior. Indicators have a social pattern or distribution but do not change their (frequency of) occurrence in different stylistic conditions. They mainly index the social characteristics of the speaker.

Table 4 The embedding of two linguistic variables

	Dentalization *Marker*	Voicing *Indicator*
Social embedding	Multi-ethnic Sharp distribution	Multi-ethnic and regional Gradual distribution
Linguistic embedding	Substrate	Structural re-setting
Interactional embedding	Yes	No

We discuss the embedding of our two linguistic variables in more detail below; starting with a synopsis in Table 5 of the results with respect to the research questions we put forth in the introduction.

Table 5 Synopsis of the results of the six research questions for the two linguistic variables

		Dentalization	Voicing
Social embedding	'regional effects'	No	Yes
	'inter-ethnic convergence effects'	Yes	Yes
	'native convergence effects'	No	No
Linguistic embedding	'structure effects'	No	Yes
	'substrate effects'	Yes	?
Interactional embedding	'style shifting effects'	Yes	No

2.6.1 Social embedding

2.6.1.1 Regional effects

To what extent are ethnolects based on local nonstandard varieties? In this study, this question is operationalized as the city of origin of the speakers, Amsterdam versus Nijmegen. No significant main effect of city on dentalization was found and this supports the conclusion that dentalization is not a feature of the traditional Amsterdam or Nijmegen dialects of Dutch – nor of any other endogenous local Dutch dialects, for that matter.

The fact that Amsterdam speakers as a group voice /z/ following vowels and sonorants significantly less often than the Nijmegen speakers do comes as no surprise in light of the fact that Nijmegen is located in a part of the Dutch language area in which voiced fricatives tend be phonetically realized as such, whereas Amsterdam is located in an area in which voicing contrasts are losing ground phonetically in the fricatives, as is the case in standard Dutch as well. So for this feature, the variation is indeed rooted in the endogenous nonstandard varieties of Dutch. It confirms the difference in index values computed in Section 2.3.1. on the basis of the sociolinguistic studies in Amsterdam (Brouwer 1989) and Nijmegen (van Hout 1989).

This regional effect is particularly strong among the 20-year-old participants. They differ between Nijmegen and Amsterdam in all three groups (cf. Figure 4). The Amsterdam speakers devoice, the Nijmegen speakers seem to voice more. Both

groups are perhaps accommodating to local regional standards. Amsterdam participates directly in the devoicing processes going on in standard Dutch in the western part of the Netherlands.

2.6.1.2 Inter-ethnic convergence effects

Do the speakers with Turkish and Moroccan backgrounds differ in their variation patterns or do they share a more global "non-native" identity? In the design of this study, this question is operationalized as the background of the speaker, a global "non-native" identity setting apart the Turkish Dutch and the Moroccan Dutch on the one hand from the endogenous Dutch on the other. As for dentalization, the Turkish Dutch and Moroccan Dutch speakers are so similar that statistically they form a group distinct from the endogenous Dutch. So, it looks as if dentalization has come to serve as an marker of a more global "non-native" identity and has hence started to develop into a "multi-ethnolect" feature, but there was evidence that the Moroccan Dutch are leading the use of dentalized variants, as they are more able to adapt dentalization depending on the interlocutor than are the Turkish Dutch. In view of the fact that – according to Dorleijn & Nortier (2006) and Nortier & Dorleijn (2008) – the Moroccan Dutch have a higher covert prestige in the Netherlands, the explanation may be socio-psychological in nature.

The very same patterning occurred for the voicing of /z/ following obstruents. Endogenous Dutch form one homogeneous subset, the Moroccan Dutch and Turkish Dutch another one. The voicing of /z/ following vowels and sonorants shows a pattern in which Turkish Dutch seem to use more voicing; cf. our considerations in Section 2.5.2.

2.6.1.3 Native convergence effects

Is there any evidence of spread of ethnic patterns of variation to endogenous Dutch peers? If so, do they occur because of "covert prestige" or because they are less marked? This question is operationalized in the first place through the background and the age group of the speakers.

Dentalization in the realization of sibilants seems to be more marked (witness, e.g. its typological frequency, viz. the fact that only 48, i.e. just over ten percent, of all languages in the UCLA Phonological Segment Inventory Database (UPSID) have dental sibilants – see Maddieson & Precoda 2011), so if this feature spreads, it is despite it being the marked option and thus probably because of extra-linguistic motives such as covert prestige. There are no significant age group effects, which may suggest that dentalization does not (yet) play a role in processes of coming of age and identity construction, but we have to be careful in drawing a conclusion as

we do not know how old the process of dentalization is. It is important to conclude that the endogenous Dutch speakers do not adopt the ethnic dentalized variants at all. These Dutch speakers seem to keep to their own track in the voicing of /z/, devoicing more than the Turkish Dutch and Moroccan Dutch speakers, and maintaining a categorical distinction between post-obstruent and post-sonorant and post-vocalic contexts.

2.6.2 Linguistic embedding

2.6.2.1 Structural effects

To what extent does endogenous Dutch linguistic conditioning (progressive assimilation in the case of devoicing, depending on the preceding context) apply in the patterns of variation found among speakers with Turkish or Moroccan backgrounds, or are conditions reset? The findings suggest that the devoicing of /z/ following obstruents (i.e. regressive assimilation of voice to a preceding obstruent) is less frequent among speakers with a Turkish or Moroccan background.

The heritage languages of such speakers and, more specifically, the fact that the literature does not contain any indications that Berber, Moroccan Arabic or Turkish have progressive voice assimilation in obstruent clusters (cf. Sections 2.3.2 and 2.3.3) seem to support this impression. It is not easy to decide whether markedness or other internal forces give rise to this development.

2.6.2.2 Substrate effects

The question that is relevant to our findings on dentalization and the voicing of /z/ is the extent to which ethnolectal variation is based on interference from the heritage languages of the speakers with Turkish and Moroccan backgrounds. This question was operationalized through the speakers' background variable in the sample.

The mean dentalization indexes appear to be highest among the Moroccan Dutch, slightly (though not significantly) lower among the Turkish Dutch, and zero among the endogenous Dutch speakers. This seems to suggest (a) that this phenomenon has its roots in the languages spoken by the Moroccan Dutch, an impression confirmed in the literature (viz. El Aissati et al. 2005), and (b) that this phenomenon has been borrowed by the Turkish Dutch. In the indigenous varieties of Dutch, dentalization is unknown and this is reflected in its total lack in the speech by the endogenous Dutch participants in our sample.

Voicing of /z/ after obstruents is virtually absent in the speech of these Dutch participants, but it occurs in about one fifth and one quarter of the realizations produced by the Moroccan Dutch and Turkish Dutch, respectively. Voicing and regressive assimilation of voicing to the preceding obstruent definitely do not have

Dutch roots in this context. Following vowels and sonorants, the endogenous Dutch as a group voice roughly one third, Moroccan Dutch almost half and the Turkish Dutch almost two thirds of tokens. In this respect, the Moroccan Dutch participants pattern with both the endogenous Dutch and the Turkish Dutch. A closer examination of the phonetic details of the original languages and acoustic measurements are needed to establish whether (in structural and/or acoustic dimensions, respectively) contrastive voicing is more important or acoustically more salient in Turkish than in Moroccan Arabic, Berber and Dutch, and perhaps different. In addition, we need to investigate if all voicing shows the same directionality in assimilation.

2.6.3 Interactional embedding

2.6.3.1 Style-shifting effects

To what extent do speakers shift between patterns of variation, depending on the background of the interlocutor? In the literature (e.g. Keim & Knöbl 2007; Kern & Selting 2011) there are indications that for speakers who control the standard or more near standard varieties, ethnolect features can offer a means to play with identities and stylization. In the design of this study, style is operationalized in connection with the interactional situation, in particular the background of the interlocutor.

The fact that dentalization indexes increase with the background of the interlocutor in line with the pattern “endogenous Dutch < Turkish Dutch < Moroccan Dutch” mirrors the degrees of dentalization in the speech of the members of these three groups. This finding is reminiscent of Bell’s (1984) insight that characteristics of the speech used in interaction with members of certain groups tends to mirror characteristics of the speech of members of these groups themselves and, more generally, that language style is a matter of audience design. The Turkish Dutch speakers have slightly higher dentalization indexes in their conversations with fellow Turkish Dutch than in those with Moroccan Dutch (let alone with endogenous Dutch) interlocutors; the Moroccan Dutch speakers have much higher dentalization indexes in their interactions with fellow Moroccan Dutch than in those with Turkish Dutch interlocutors. Speakers from both groups show the lowest dentalization indexes in their conversations with endogenous Dutch and the highest indexes in conversations with speakers of the same background, which suggests that dentalization is an in-group feature for both groups. In this respect, the behavior of the Moroccan Dutch is even more pronounced than that of the Turkish Dutch. Compare Fagyal & Stewart’s (2011:75) finding of an intonational variant which “seems to function as a micro-level style feature indexing common ground” among preadolescents in a Paris banlieu.

Breaking down the dentalization picture for the two cities, it appears that the

Moroccan Dutch behavior is fairly consistent, whereas the Turkish Dutch in Amsterdam behave differently from those in Nijmegen. While the Nijmegen Turkish Dutch seem to handle dentalization as an in-group feature (with their dentalization indexes with Moroccan Dutch interlocutors being barely higher than those for endogenous Dutch interlocutors), the Amsterdam Turkish Dutch have relatively high dentalization indexes for both endogenous Dutch and Moroccan Dutch interlocutors and low ones for fellow Turkish Dutch; dentalization may act as a *trait d'union* with the Moroccan Dutch (for whom dentalization is endogenous) but, on the other hand, as a way to diverge from the endogenous Dutch. The differential patterning of dentalization for both groups suggests that it is a relatively stable and probably older in-group feature for the Moroccan Dutch, but a more recently acquired identity-marking resource for the Turkish Dutch, given the difference between the Amsterdam and Nijmegen Turkish Dutch in this respect.

Voicing appears to be different, as no style-shifting effects were found. Stronger voicing is an indicator of ethnicity, but a gradual, quantitative one, with the strongest differences in the postobstruent context. Even in this context, no style shift was observed, meaning that a structural resetting of the linguistic conditions by the Turkish and Moroccan Dutch goes unnoticed as a potential resource for stylization in interaction.

The structural resetting in combination with the absence of a stylistic shift in voicing makes clear that the language of the Turkish Dutch and Moroccan Dutch is better defined as (multi-)ethnolect than as youth language. It shows the structural nature of voicing in their grammars. It seems to have developed as a stable, long-lasting feature, which may even affect the status of progressive assimilation in Dutch over time.

2.7 Looking ahead

Further research is needed to establish whether contrastive voicing is more important in Turkish than in Moroccan Arabic, Berber and Dutch,[5] and to explain the apparent social spread of /z/ voicing following obstruents and regressive voice assimilation. Another desideratum would be to include more linguistic variables, to investigate their social and ethnic distribution. Other phenomena to be studied include the realization of the diphthong /ɛi/ and the tenseness/laxness of long vowels in closed syllables.

In the ideal case, the findings could also be compared to findings for similar groups in similar cities in Germany, Sweden, England, etc. Obviously, international research cooperation (starting at the level of the Germanic language family for

[5] This is despite the fact that /z/ is relatively marginal in Turkish Dutch; cf. Section 2.3.2.

instance) could greatly help to unravel internal, external and extralinguistic forces underlying the development and use of new nonstandard features as well as their diffusion to the verbal repertoires of other cultural groups.

3 Variation in the realization of /ɛi/ by Dutch youngsters: from local urban dialects to emerging ethnolects?

Edited from: van Meel, Linda, Frans Hinskens & Roeland van Hout. 2014. Variation in the realization of /ɛi/ by Dutch youngsters: from local urban dialects to emerging ethnolects? *Dialectologia et Geolinguistica* 22. 46–74. doi:10.1515/dialect-2014-0004

Abstract

How do speakers of current Turkish and Moroccan ethnolects of Dutch deal with phonemes that do not exist in their heritage languages and that are at the same time subject to pronounced regional and social variation in the Dutch speech community at large, such as the Dutch diphthong /ɛi/? This diphthong does not occur in Turkish and Berber and it occurs only as a dialectal allophone in certain dialects of Moroccan Arabic.

Data from speakers from the Amsterdam and Nijmegen urban areas are studied. In the Amsterdam dialect, the diphthong is traditionally subject to monophthongization and lowering, yielding realizations as [æ:] or [a:], while in Nijmegen the diphthong is only subject to monophthongization, resulting in the variant [ɛ:]. Recently, a new lowered, diphthongal variant [ai] entered colloquial spoken standard Dutch. Therefore, ethnolect speakers have a wide range of variants to 'choose' from: the traditional standard Dutch variant [ɛi], the new variant [ai], which is expanding areally and socially, and the monophthongal variants of the surrounding urban dialects.

Two variable properties of /ɛi/ are examined: (1) height of the prominent first element, and (2) the degree of monophthongization. The urban dialect features which had developed into sociolect features over the past generations appear to be undergoing social redistribution to become ethnolect markers.

3.1 Introduction

Ethnolects are a new domain in the study of language contact and bilingualism, where they have so far mainly been looked at from an ethnographic angle (see Section 3.2 below). The present contribution focuses on analyzing patterns of linguistic variation from a language centered, sociolinguistic perspective using quantitative tools to analyze patterns of linguistic variation.

The past and present-day emergence of Dutch ethnolects is detailed in Hinskens (2011) and Muysken (2013); among the historical ethnolects is Jewish Dutch, while Surinamese Dutch is one of the modern ethnolects. In this contribution, we present findings from a recent research project based on data from speakers of Moroccan and Turkish ethnolectal varieties of Dutch living in the cities of Amsterdam and Nijmegen. These ethnolectal varieties of Dutch originated in labour migration which occurred in the 1970s (Hinskens 2011; Muysken 2013). Moroccan-Dutch and Turkish-Dutch are, therefore, 'immigration varieties'. Both ethnolectal varieties have

been investigated by several researchers, with Dorleijn & Nortier (2006), El Aissati, Boumans, Cornips, Dorleijn & Nortier (2005), Nortier & Dorleijn (2008), van Krieken (2005), van Lier (2005) all giving overviews of some features of Moroccan-Dutch and Turkish-Dutch. Other researchers have investigated these new ethnolects in relation to the notions of stylization and identity (cf. Flanders: Jaspers 2006, 2008; Jaspers & Aertsen 2004; Netherlands: Nortier & Dorleijn 2008).

In previous research, we reported a dental, voiced /z/ as a characteristic of the Dutch ethnolects of Turkish-Dutch and Moroccan-Dutch speakers in the cities of Nijmegen and Amsterdam (van Meel, Hinskens & van Hout 2013). This dental realization has its origin in the Moroccan languages and is not part of the dialectological and/or sociolinguistic patterns of variation of traditional endogenous Dutch.

In the present study, we use the same corpus to investigate how speakers of current Dutch ethnolects deal with Dutch phonemes which do not occur in the heritage languages involved, but are marked by intricate regional and social stratification patterns in varieties of Dutch. A phoneme that unequivocally meets these criteria is the Dutch diphthong /ɛi/ that belongs to the set of three diphthongs of modern Dutch, all of which are mid-high closing, the first element being the prominent one. Such a diphthong does not occur in Turkish and Berber and does not belong to the phoneme inventory of the Moroccan Arabic dialects either (see Sections 3.3.2 and 3.3.3 below). Therefore, its absence from the heritage languages involved seems to preclude substrate effects.

The Dutch diphthong /ɛi/ can be pronounced in varying ways. A first relevant distinction is the one between the urban dialects of Amsterdam and Nijmegen (i.e. the two cities of our study). In Amsterdam, this diphthong is subject to monophthongization and lowering, leading to realizations as [æ:] or [a:] (Brouwer 1989; Schatz 1986). In Nijmegen, the diphthong is just subject to monophthongization, leading to the variant [ɛ:] (van Hout 1989, 1999); both variants have low overt (and no covert) prestige. A second relevant distinction concerns the spoken standard variety. From the early 90's onwards, a new variant of the diphthong /ɛi/ has been observed in colloquial standard Dutch: the lowered, diphthongal variant [ai]. This change in the colloquial standard Dutch pronunciation of /ɛi/ has been claimed to have been propelled by well-educated young women and is seen by some linguists as typical of a new, relatively informal standard Dutch variety called 'Polder Dutch' ("Poldernederlands", cf. Stroop 1998). This variant also serves as the perfect 'compromise' between traditional dialect variants (which are subject to generally ongoing processes of dialect leveling – Hinskens, Auer & Kerswill 2005), on the one hand, and the variant which is in line with the standard norms, on the other. All this

means that ethnolect speakers have a pool of variants they can 'choose' from: the traditional standard Dutch variant [ɛi], the new, expanding variant [ai], and the local monophthong variants marking the dialects of Amsterdam and Nijmegen.

By the end of the 19th century, standard Dutch had started to take root in oral usage in the higher status groups. As a result, especially urban dialects increasingly became sociolects, marked by a socially stratified linguistic continuum between the urban dialect and standard Dutch. Nowadays, urban dialects are spoken in their most pronounced form in the low-income neighborhoods of Nijmegen and Amsterdam (cf. Brouwer 1989; Schatz 1986 for Amsterdam; van Hout 1989, 1999 for Nijmegen) and their prestige is relatively low.

Former immigrants from Turkey and Morocco and their families typically live in densely populated neighborhoods with cheaper housing, lower incomes, higher unemployment rates and reduced access to infrastructure. Growing up in these areas situated at the lower ends of the socio-economic continuum brings the second and third generation immigrants into contact with peers of Dutch descent who use urban dialect as their native speech in the neighborhood as well as in school. This situation of long-lasting intensive exposure seems to be an ideal context for youngsters with a different ethnic background to converge to the surrounding local urban dialect. If they did not simultaneously develop or adopt additional ethnic linguistic markers, they would become undistinguishable from 'white' speakers of the local urban dialects. However, if at the same time their white peers were decreasing the distance to the standard language by using more standard language features, then local markers would become ethnic markers. The end effect would be a social-cultural redistribution of variants: urban accent variants would become ethnic markers.

The main question we will try to answer is whether /ɛi/ is involved in processes of local redistribution. We will test the assumption that the variation patterns in the realization of the diphthong will have their main origin in the local urban dialects of Nijmegen and Amsterdam. Since these local urban dialects are socially stratified plus the norms for spoken colloquial standard Dutch are changing, we *hypothesize* that

1. white Dutch boys will target the more prestigious standard forms of standard Dutch (local divergence, upward convergence toward the standard norm),
2. leaving the traditional dialect to the ethnic groups (local convergence towards the socially low prestige urban dialect). Urban dialect variants become part of the ethnolects. This would amount to a two-step resetting of the social distribution of the variation in the realization of /ɛi/. Local dialect variants would, thus, change from sociolectal to ethnic markers. Since there is no reason to expect any substrate effects with respect to the realization

of this diphthong, we hypothesize

3. that there will be no differences between both ethnic groups, resulting in a multi-ethnolect feature (inter-ethnic convergence).

The design of our database (which will be presented in Sections 3.4.1 and 3.4.2 below) enables us to test our hypotheses, taking into account the role of linguistic conditions (the linguistic embedding, including substrate effects), and the way variation may play an active role in face-to-face interaction (style shifting dependent on the interactional embedding, determined by the background of the interlocutor).

In this contribution, we first will zoom in on the ethnolect concept and some of the main recent studies (Section 3.2), the different realizations of /ɛi/ and sketches of the phonetic and phonological characteristics of the diphthong /ɛi/, of its variants in the Amsterdam and Nijmegen dialects, and its nearest neighbors in the languages at issue (Section 3.3). The methods are discussed in Section 3.4, followed by a presentation of the main results in Section 3.5. Section 3.6 contains the discussion and conclusions.

3.2 Ethnolect: concept and scholarship

There has been heated debate on the definition of the notion of ethnolect and its relation to concepts such as youth language and nativized varieties. The notion of ethnolect was introduced in the late 1970's to refer to "the English of the descendants of immigrant families long after their original language is lost" (Carlock & Wölck 1981; via Wölck 2002:157). Danesi (1985:118) has defined ethnolect as "the variety of a language that results when speakers of different ethnolinguistic backgrounds attempt to speak the dominant language (e.g. Chicano English)". Unlike dialects, koines (or 'regiolects'), and other homegrown varieties of a language, and unlike most transplanted varieties[1] ethnolects typically are not the mother tongue of the first generations of speakers. Unlike transplanted varieties and 'daughter languages', as products of language shift, ethnolects commonly develop in the language area or at least in the interaction with native speakers of the language. Ethnolects are not necessarily learners' varieties as many speakers have a command of the standard variety. For such speakers, it is not a matter of not being able to speak the standard, but rather of not wishing to speak the standard variety in certain domains or settings.

Most definitions of the ethnolect concept are stipulative (rather than descriptive); they do not describe empirically established distinctive features, rather, they are conventions on what is meant by that notion. Clyne (2000:86), for instance, has

[1] I.e. varieties of a language which are spoken in a community outside the original speech community, such as the descendants of Dutch dialects spoken in Iowa (Smits 1996).

defined ethnolects as "varieties of a language that mark speakers as members of ethnic groups who originally used another language or distinctive variety". Androutsopoulos (2001:2) defined an ethnolect as "a variety of the majority language (or 'host language') which is used and regarded as a vernacular for speakers of a particular ethnic descent and is marked by certain contact phenomena". According to Auer (2003:256), "an ethnolect is a way of speaking (a style), which by the speakers themselves or by others is associated with one or more non-German ethnic groups" [our translation]. In Auer's conception, an ethnolect (also) concerns grammar, as opposed to the known innovations of youth language, which does not. Muysken (2013) describes more such oppositions; in his view, ethnolects are more or less stable and their usage is at most semi-conscious; ethnicity plays an inherent role; and the features involved are phonology and syntax. In Muysken's view, youth language is dynamic and its usage is (semi-) conscious; ethnicity plays a role only temporarily and the features are usually lexical and pragmatic in nature (cf. Auer 2003: 256). Youth language[2], thus, seems to be more of a register or even jargon.

With regard to the functional dimension, the question arises as to whether ethnolects are Mediums for Inter-ethnic Communication or rather Mediums for Community Solidarity, in Baker's (2000) terminology. In the latter case, ethnolects will probably function mainly or merely as in-group codes; in that case, the emblematic value of the ethnolectal variants, which are often quite distinct from the prestigious norm, is mainly defined by their signaling ethnic identification and solidarity. This is line with Benor's (2010:160) concept of ethnolinguistic repertoire "as a fluid set of linguistic resources that members of an ethnic group may use variably as they index their ethnic identities".

Features which originated in the language contact situation underlying the development of a specific ethnic variety sometimes spread to other ethnic groups to become (what has been referred to as) 'multi-ethnolect' features (Clyne 2000; Quist 2000; Wiese 2009, 2013). Multi-ethnolect features can also be stabilized second language acquisition phenomena and they, thus, need not be specific to any ethnic group. An example is variation in the marking of Dutch grammatical gender (Hinskens et al. in preparation).

Generally, two approaches to the study of ethnolectal variation can be distinguished: the language centered and the ethnographic approach. Whereas the ethnographic approach conceives language systems as infinite resources from which

[2] Dutch: jeugdtaal. The late modern urban manifestations are also referred to as straattaal, lit. 'street language'; cf. van Lier (2005).

speakers may freely[3] choose to construct their identity, the language centered approach tries to disentangle the laws, generalizations and restrictions on these resources, with a distinct emphasis on linguistic embedding. Terminology such as 'ethnolect', 'multi-ethnolect' and 'multicultural variety' is characteristic of the language centered approach. This approach, which highlights features of linguistic structure, their origin and distribution, is quantitative – often in the Labovian tradition. The features' patterns of use are usually viewed from a rather macro-social angle (e.g. Hoffman & Walker 2010; Cheshire et al. 2011). The ethnographic approach, in contrast, which stands out by terminology such as 'style' and '(pan-) ethnic style' (see Kern 2011b:9 and the reference cited there), is typically couched within the Gumperzian concept of style as a comprehensive and 'weak' concept. In this view, styles have a prototype organization. Accordingly, 'ethnic ways of speaking' are seen as "rather fuzzy phenomena with some core linguistic features and more variable features at the boundaries" (Kern 2011b:9). Styles are not determined; they are continuously being constructed as socially and interactively significant products (cf. Eckert's (2008b) view that ethnolects are fluid rather than fixed entities). Both reactive and initiative uses of linguistic and paralinguistic features are analyzed; in the latter case, speakers proactively select from various linguistic resources – possibly for the purpose of changing the situation or presenting themselves in a certain way. The angle is micro-social and interpretive (e.g. Benor 2010; Keim & Knöbl 2007; Kern & Selting 2011). In contrast to the language centered approach, in the ethnographic approach language change is not a central concern.

What is also illuminating is the diverging view on ethnicity characterizing both approaches. Whereas the language centered approach tends to employ an 'etic', 'objective' definition of ethnicity (operationalized through variables such as language, race, and descent), the ethnographic approach typically applies an 'emic', 'subjective' definition of ethnicity as a social construction, in which perception plays a crucial role as well.

3.3 The /ɛi/ in the languages involved

Some of the earlier mentioned studies discuss the diphthong /ɛi/. According to El Aissati et al. (2005:162), members of the first generation of Moroccan and Turkish speakers realise the sound as [e:i], [ɑi] or [a:i], while the second generation speakers in their study used the plain standard Dutch [ɛi]. Referring to El Aissati et al. (2005), Nortier & Dorleijn (2008) discuss the language use of Moroccan Dutch of the second generation; the authors point out that "their pronunciation of tense vowels and

[3] Linguistically free, i.e. without discernable internal conditioning – as in the structuralist concept of 'free variation'.

diphthongs is consistent and indistinguishable from native speakers" (p.130). Van Krieken (2004:62) carried out a pilot study of variation in the realization of /ɛi/ in Moroccan Dutch in Nijmegen. Only 11.8% of the cases were monophthongized, i.e. [ɛ:] (cf. the sketch of the dialect of Nijmegen in Section 3.3.1.3), and she did not find any style effects for the variable /ɛi/. In this section, we overview the /ɛi/ in the languages involved.

3.3.1 Dutch

3.3.1.1 Standard Dutch

Dutch has one reduced vowel (/ə/) and fifteen full vowels (excluding loan phonemes) which can be subdivided into five lax vowels (/ɛ, ɪ, ʏ, ɑ, ɔ/), seven tense vowels (/i, y, a, u, e, o, ø/) and three diphthongs (/ɛi, œy, ɔu/) (Booij 1995; Gussenhoven 1999). The three diphthongs can be defined as sequences of two non-identical vowels (Booij 1995). The two vowels within a diphthong only differ in height. The first vowel has the features [-high] and [+mid], and the second one [+high] and [-mid] (Booij 1995). Graphic representations can be seen in Gussenhoven (1999:76) and (based on acoustic measurements) in Van der Harst (2011:e.g. 328). Phonologically, the two elements of the diphthong are identical with respect to [back] and [round], i.e. the diphthong /ɛi/ is [-back] and [-round]. The diphthongs must be distinguished from vowel + glide combinations such as /ɑj/ that occur in a few Dutch words such as *mais* /mɑjs/ 'corn' (Booij 1995).

In Dutch, the diphthong /ɛi/ can be represented orthographically in two ways: <ei> and <ij>, since two etymologically and phonologically distinct historical vowels have merged in the dialects which were eventually promoted to standard Dutch. While the written forms may have distinct lexical meanings, as in *reizen* 'to travel' and *rijzen* 'to rise', the pronunciation is homophonous, i.e. [rɛizə]. The distinction in writing is due to the fact that in an older historical phase, <ei> was pronounced as [ai] and <ij> as [i:] (cf. Spiegel 1962 [1584]:20, 26). From the second half of the 16th century onwards, a standardization process took place in the Low Countries (cf. van der Wal 1992) and the two sounds merged, resulting in Standard Dutch [ɛi]. In other words, the distinction between <ei> and <ij> is etymological and orthographic. However, in many Dutch dialects the distinction between <ei> and <ij> is still phonological (cf., e.g. maps no. 56, 57, 60, 61 of the "Fonologische Atlas van de Nederlandse Dialecten (FAND)" [Goossens, Taeldeman & Verleyen 2000]).

As noted in Section 3.1, Stroop (1998) noticed a lowered variant of the diphthong /ɛi/ pronounced mainly by well-educated young women in standard Dutch.

He dubbed the standard language variety spoken by these women 'Polder Dutch'.[4] This process of a lowering diphthong had already happened in the languages of our neighbors: English and German as well as in a subset of Hollandic dialects. These languages have a diphthong starting with a low front vowel [aɪ], i.e. compare English <wine> [waɪn] and German <Wein> [waɪn] with Dutch <wein> [wɛin].

Applying acoustic analyses, Jacobi (2009) examined the Dutch diphthong /ɛi/ in the Spoken Dutch Corpus (Oostdijk et al. 2002), a large corpus of standard Dutch in which many speakers with different backgrounds participated. She found that there was no difference between men and women in her study, and she therefore rejected Stroop's hypothesis that women are leading (however, see van Heuven, van Bezooijen & Edelman 2005 for a finding which is line with Stroop's claims). Jacobi did find differences between social classes. Speakers with a strong educational and occupational background were found to have both a lower onset and stronger diphthongization than speakers with a more limited educational background.

3.3.1.2 The local dialect of Amsterdam

The Amsterdam vernacular is characterized by (a) lowering of the prominent first element of the diphthong and (b) monophthongization. Example (1a) shows an example of pure monophthongization. In examples (1b) and (1c), lowering took place in addition to the monophthongization; in addition (1c) shows retraction.

(1) Examples of the Dutch word [pɛin] 'pain' pronounced in the dialect of Amsterdam (Brouwer 1989:29–30; Schatz 1986:65):

a) /ɛi/ → [ɛ:] [pɛ:n]
b) /ɛi/ → [æ:] [pæ:n]
c) /ɛi/ → [ɑ:] [pɑ:n]

3.3.1.3. The local dialect of Nijmegen

In the Nijmegen dialect, /ɛi/ used to be pronounced as /i(:)/. Nowadays, however, this old variant hardly occurs in spontaneous speech anymore[5] (van Hout 1989:86–87). Instead, the standard Dutch variant [ɛi] is used, as well as a monophthongized variant (see 2b).

(2) Examples of the Dutch word [pɛin] 'pain' pronounced in the dialect of Nijmegen:

a) /ɛi/ → [ɛi] [pɛin]
b) /ɛi/ → [ɛ:] [pɛ:n]

[4] More on Polder Dutch on Stroop's website: http://cf.hum.uva.nl/poldernederlands/english/main_engels.htm, http://cf.hum.uva.nl/poldernederlands/index.html

[5] In reading tasks, the old variant [i(:)] certainly is no longer used (van Hout 1989).

3.3.2 Turkish

Turkish has eight phonemic vowels: /a, ɯ/ɪ, o, u, e, i, ø/œ, y/. Turkish does not have diphthongs with phonemic status (cf. Comrie 1997; Göksel & Kerslake 2005; Kornfilt 1997; Lewis 2000[1967]). Swift (1963:11) states that "Turkish has a few diphthongs", but he does not include them in the list of phonemes. This statement seems to be grounded in the observation that "most vowels may be followed by /y/ [i.e. [ɪ], LvM] with a diphthongal result much like the diphthong sound of English say".

3.3.3 Moroccan languages

Morocco is a multilingual country. Both Moroccan Arabic and Berber languages are used as mother tongue(s) by Moroccans (cf. El-Aissati et al. 2005:150; El-Aissati & E-rramdani 2001:63).

Moroccan Arabic consists of several dialects which are mutually intelligible. The Berber languages can be divided into three main groups: 1. Tarifit, 2. Tachelhit (also written as Tachelhiyt or Tashelhiyt), 3. Tamazight (Central Atlas). These three language groups are not mutually intelligible.

According to a rough estimation by El Aissati et al. (2005:150), some 60 percent of Moroccans in the Netherlands speak Tarifit, 10 percent speak Tachelhit and the remaining 30 percent speak a Moroccan Arabic dialect as their mother tongues.

In the next two sections, the vocalic phonemes as well as the diphthong /ɛi/ are discussed for Moroccan Arabic and for the Berber languages Tarifit and Tachelhit.

3.3.3.1 Moroccan Arabic

The dialects of Moroccan Arabic have three full vowels in common: /i/, /u/ and /a/, which are considered to be vocalic phonemes. Besides these phonemes, there are many vowel allophones which can vary among dialects (Abdel-Massih 1973:23 lists about 12 of them).

Moroccan Arabic also has diphthongs or diphthongal elements, but the dialects differ in their realization and structural status. A diphthong in one dialect can correspond to a monophthongized element in another; e.g. different forms for 'eggs' (Heath 2002:199) are *bayṭ* (northern, Jebli), *bəyḍ ~ bǎyḍ* (rural belt, Atlantic strip), *biḍ* (urban belt), and *bǎyð* [bɛːð] (Saharan dialects). If a dialect has the diphthongal element *ay*, it depends on the dialect whether it has phonological status (Heath 2002:197). In some dialects, the short diphthongs {ey ew} and {ǎy ǎw} are often pronounced as [eː], [oː], [ɛː] or [ɔː], but, according to Heath (2002:198), "a diphthongal phonemic representation {ǎy ǎw} still seems appropriate". Other dialects with diphthongal elements have long diphthongs with a full /a/ vowel {ay aw}. In other words, if dialects have diphthongal elements, these elements can be either short, i.e.

ey ew ăy ăw, or long, i.e. *ay aw*. These six diphthongal elements are also mentioned by Harrell (1962:14–15, 1965:8) who describes their pronunciation. According to Abdel-Massih (1973:31), Moroccan Arabic has five diphthongs: *aw, ay, iy, iw, uw.*

3.3.3.2 Berber

The dialect variation in the Berber phonological systems is large. Just like Moroccan Arabic, most Berber dialects have three vocalic phonemes: /i/, /u/ and /a/ (Kossmann & Stroomer 1997:463). Lafkioui (2007:17) also mentions these three 'voyelles de base' for the varieties of Tarifit (i.e. 'variétés berbères du Rif'): /i/ is pronounced basically as [ɪ] and in certain conditions as [ɪˑ], [e] or [eˑ]; /u/ is basically realized as [ʊ] and in specific contexts as [ʊˑ] or [o]; /a/ is basically pronounced as [æ] or [ɛ] and in particular environments as [æˑ], [ɛˑ], [a], [ʌ] or [ʌˑ]. McClelland III (2008) also states that there are three vocalic phonemes in Tarifit, but instead of /a/, he claims the third vocalic phoneme is /æ/ (i.e. the three are /i/, /u/ and /æ/). He lists ten 'vocalic phones' that are known for this dialect group (McClelland III 2008:26).

As regards diphthongs, Lafkioui (2007) and McClelland (2008) do not list any, and Kossmann & Stroomer (1997) explicitly claim that there are no diphthongs (and that Berber does not have vowel clusters either).

3.4 Methods

3.4.1 Participants

Data were collected from 51 youngsters from three different groups: Moroccan-Dutch, Turkish-Dutch and white Dutch. The groups were controlled for age (10–12 versus 18–20 years old), residence (Amsterdam vs. Nijmegen) and, in the case of the white Dutch speakers, the presence or absence of regular contacts with Turkish-Dutch and Moroccan-Dutch friends. Table 1 presents the speaker design.

Table 1 Overview of the research design and the number of participants

Background	Moroccan-Dutch		Turkish-Dutch		white Dutch			
Inter-ethnic ties?	Yes		Yes		Yes		No	
years of age	10-12	18-20	10-12	18-20	10-12	18-20	10-12	18-20
Amsterdam	3	4	3	4	2	3	3	3
Nijmegen	3	3	3	3	3	3	4	4

The participants with a Moroccan or Turkish background were born in the Netherlands and have at least one parent who immigrated from Morocco or Turkey respectively. All participants grew up in their place of residence (either Amsterdam or Nijmegen). The Dutch group was separated into two groups, those with ethnic ties and those without, that is those who have friends in their social networks with a Moroccan or Turkish background or not. We will refer to the participants with

Moroccan and Turkish backgrounds as Moroccan-Dutch and Turkish-Dutch. The two groups with an entirely Dutch background will be referred to as white Dutch, i.e. either with (D-group) or without inter-ethnic ties (C-group).

Due to unanticipated complications during the fieldwork sessions, there is only scant information available about the language skills and linguistic profiles of the speakers in our sample. The information available is detailed in van Meel, Hinskens and van Hout (2013). In short, all Turkish-Dutch speakers reported knowledge of Turkish and all Moroccan-Dutch speakers reported knowing either one or more of the Moroccan languages. No reliable information regarding the speakers' relative proficiencies in the languages mentioned is available – nor in the varieties of the languages mentioned.

3.4.2 Material and data

The participants had free conversations of about one hour with a partner. Each conversation involved two peers of the same age group (either two 10–12 year olds or two 18–20 year olds).

The speakers with a white Dutch background with inter-ethnic ties as well as those with a Moroccan-Dutch or Turkish-Dutch background were recorded in at least one in-group conversation and two different out-group conversations. In an in-group situation, the conversation partner was a peer from their own ethnic group, while in the out-group situations, they talked with a peer from the other two ethnic groups. The recordings last about one hour for each pairing, so for each of these speakers in total approximately three hours of conversational speech were recorded.

The white Dutch participants with no inter-ethnic ties (C-group) were merely recorded in in-group conversations with fellow Dutch participants that equally have no inter-ethnic ties. They serve as the control group for the white Dutch participants with strong inter-ethnic ties (D-group).

The majority of the speakers attended the same school as their conversation partners, and many of them were classmates (especially so in the case of the 10–12 year olds). The interviewer was only present at the beginning and the end of the conversation to ensure a more natural conversation. However, especially so in the case of the 10–12 year olds, additional guidance was needed to keep them talking for one hour (i.e. suggesting topics to talk about, introducing card games).

The conversations were recorded on a Marantz Professional CD recorder CDR300. Ten to 15 minutes of each recording was transcribed using the multimedia annotator Elan (cf. Brugman & Russel 2004). These transcriptions were checked by a second transcriber.

The first three minutes of each conversation were skipped to give the

participants some time to get used to being taped and pay less attention to their way of speaking. We defined a set of criteria to select proper /ɛi/ words, excluding variation (e.g. shortening) that was not relevant in investigating patterns of monophthongization and lowering. The /ɛi/ words had to meet the following criteria:

1. /ɛi/ must be followed by a plosive or a fricative.
2. /ɛi/ must have either primary or secondary stress.
3. No /ɛi/ from the suffix *-heid* was selected as this is a frequent derivational affix which never has main stress.
4. Words with an orthographical <ij> or <ei> which according to the standard norms are not realized with [ɛi] were excluded (e.g. *bijzonder* /bizɔndər/ 'special', *vriendelijk* /vri:ndələk/ 'friendly').
5. Reduced realizations were excluded as well (e.g. *zijn* 'are/to be' reduced to *z'n* [zən], *altijd* 'always' reduced to alt').
6. To ensure the data were not biased by a specific high frequency word, a word was selected at most three times for a given speaker in a given conversation.
7. Words which were read from, for example, (news)papers and magazines were not taken into account.
8. Words which were uttered in an (intentionally) conspicuous way were not selected either. For the most part, these were imitations.

From each conversation, 6 to 13 /ɛi/-words per speaker were selected, with a mean of 9,89.

3.4.3 Variants and coding

The transcription and coding of the realization of /ɛi/ was done by the first author. The realizations were rated for height of the first, prominent element and for the degree of monophthongization. The height of the first, prominent element was coded on a four-point scale and the degree of monophthongization on a three-point scale.[6] Both scales are shown in Table 2. An /ɛi/ pronounced according to traditional Standard Dutch norms, i.e. [ɛi], (see Section 3.3.1.1) was coded as height 2 and monophthongization 1.

[6] Originally, we had a four-point scale, distinguishing shortened and normal monophthongized variants. In the analyses, we excluded the shortened forms.

Table 2 Coding schemes for (a) height and (b) monophthongization

(a) Height			(b) Monophthongization	
1	e or ɪ	(mid-)high / close	1	No monophthongization (Clear diphthong)
2	ɛ	mid	2	Slightly monophthongized (Minor diphthong)
3	æ or ɑ	mid-low / near open	3	Monophthongized
4	a	low / open		

The reliability of the coding was tested by having 261 realizations coded by a second rater, a trained phonetician, who did not know the previous ratings. The realizations were coded by the second rater in their original context. These realizations were selected from the in-group data of the 20-year-olds. Realizations by younger speakers were not analyzed, as their higher voices may cause problems in carrying out acoustic analyses. This resulted in six to twenty /ɛi/s per speaker. The inter-rater reliability was checked by calculating Cronbach's alpha. The reliability value was high (.888) for height and had an even higher value (.951) for monophthongization.

Acoustical measurements were carried out. To this end, two realizations of /ɛi/ per speaker were selected from the realizations that had been checked by the second rater. The selected realizations had to meet the criterion of being scored identically by the first and the second rater. From each speaker, two tokens of the vowel /a/ and two tokens of the vowel /i/ were also selected. The word in which a token occurred was either a noun, a verb or an adjective, and the vowel (a) was followed by an obstruent, (b) occurred in a stressed syllable, and (c) was not reduced audibly. Six vowels per speaker were measured using the default settings for measuring formants in Praat (Boersma & Weenink 2010).[7] The vowel /ɛi/ was measured at 25 and 75 percent of the duration of the vowel (cf. van der Harst 2011). The vowels /a/ and /i/ were measured at 50 percent of the duration. The outcomes of the analysis of Praat were checked manually. Pearson correlations showed a strong correlation of .801 between the F1 at 25 percent of the duration and the scores on height. They showed a correlation of .802 as well between the acoustical measure of monophthongization and the expert scores on monophthongization. This relative measure was computed by taking the absolute diphthongization of the /ɛi/-realization, i.e. the F1 at 25% minus the F1 at 75%, and dividing this by the difference of the minimum and the maximum value of F1 of the vowels (at 25% of the duration of /ɛi/ and at 50% of /a/ and /i/); a score of 1 means a consistent use of diphthongs, etc. The acoustic measurements underlined the validity of the rater scores.

[7] In these settings, at first the frequencies above 50 Hz are pre-emphasized, resulting in an amplification of +6 dB for the frequencies around 100 Hz, +12 dB for the frequencies around 200 Hz, and so on. Thereafter the Burg-algorithm is applied to a Gaussian window with a length of 25 ms, which shifts every 10 ms, to obtain the actual formant estimation. The cut-off frequency was set to 5000 Hz and the number of LPC coefficients used is 10.

In addition to height and monophthongization, the /ɛi/ tokens were coded as well for several linguistic conditions: (1) IJ-EI – i.e. whether /ɛi/ is etymologically/orthographically <ij> or <ei>; (2) word class – i.e. whether the word was a content word or a function word (pronouns and prepositions); (3) open vs. closed syllable – i.e. whether /ɛi/ was in an open or closed syllable; (4) LOG frequency – i.e. the log of the frequency of the citation forms. This latter measure was calculated by adding up the frequencies of all related tokens in the spontaneous conversations and telephone dialogs of the Spoken Dutch Corpus (Dutch Language Union 2004) that belong to the same citation form (lemma) and then applying a logarithmic transformation to the frequency of the citation form.

3.4.4 Data analyses

The data for the auditory ratings concerning height and monophthongization of /ɛi/ were analyzed in order to establish the effects of the factors we defined, in particular the four social factors. We start the analysis with an overall analysis of the performance of the four speaker groups, including the age and city distinctions. The mean value over all tokens was computed per participant. The results will be presented in histograms, followed by an analysis of variance, to evaluate the impact of age, city and the background of the speaker. In the next step, we turn to the impact of the background of the interlocutor. In presenting the results, we first focus on the three groups that interacted with three different interlocutors, excluding the C group (white speakers with no inter-ethnic ties). We will present figures to gain a good impression of the outcomes. The statistical analysis carried out was a repeated measures analysis of variance.

In the final statistical analysis, we included all groups and all factors, applying a mixed model analysis, starting with the social factors. The four linguistic factors as described in Section 3.4 were included indirectly in this analysis, by controlling them by using citation form (lexical entry) as a random factor. The impact they have will be discussed in the final analysis, where they are added as separate factors. Implementing the linguistic factors will turn out not to influence the role of the social factors. The outcomes for monophthongization will be presented in Section 3.5.1, those for height in Section 3.5.2.

Section 3.5.3 combines the dimensions of monophthongization and height in a multidimensional scaling analysis which was carried out in order to uncover the position of speakers relative to each other. The aim is to gain an insight into the distribution of the speakers in recognizable subgroupings.

3.5 Results

3.5.1 Monophthongization

Figure 1 summarizes the outcomes of monophthongization for the two cities, Amsterdam versus Nijmegen, split up for background of the speaker and age.

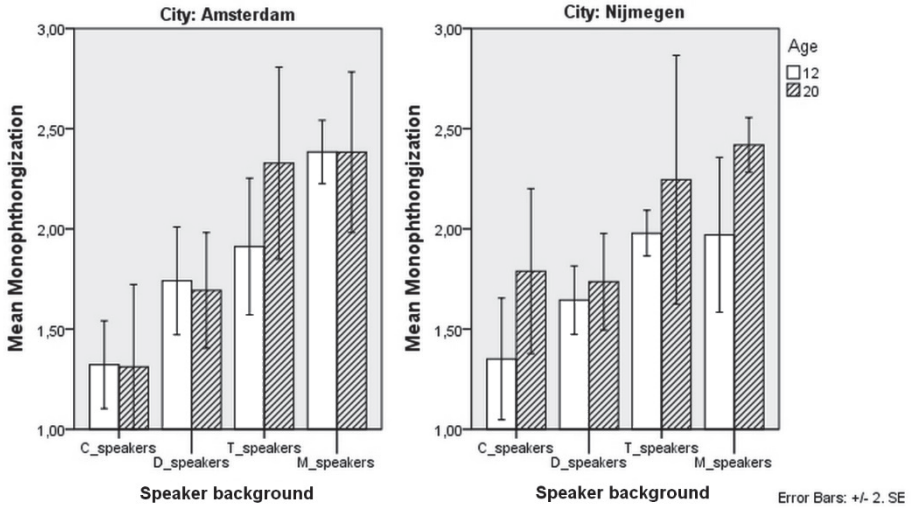


Figure 1 The mean monophthongization scores in Amsterdam (left) and Nijmegen (right), split up for the background of the speaker and the two age groups; the higher the values the more monophthongization. The C and D speakers are the white Dutch speakers with no inter-ethnic ties (C) and strong inter-ethnic ties (D). The T speakers are Turkish-Dutch; the M speakers Moroccan-Dutch. The speaker variation within the groups is indicated by the error bars

Figure 1 shows that the differences between the cities are small overall. In both cities, the white Dutch groups have the lowest scores. No differences between the Turkish-Dutch and Moroccan-Dutch group seem to exist.

The ANOVA results confirm our impression. There is no main effect for city ($F < 1$) and a strong effect for the background of the speaker ($F(3,35) = 18.617$, $p=.000$, partial eta squared = .615). There is a moderate overall age effect as well ($F(1,35) = 4.789$, $p=.035$, partial eta squared = .120), indicating higher monophthongization scores for the 20-year-olds. None of the four interaction effects was significant.

The effect for speaker background was investigated in more detail by applying a post hoc analysis (Tukey HSD). The post-hoc analysis makes plain that statistically there are but two groups: white vs. 'non-white'. Speakers with a Dutch background have a mean monophthongization value of 1.46 (no ethnic ties) and 1.70 (strong ethnic ties), respectively, while the Turkish-Dutch and the Moroccan-Dutch speakers have 2.13 and 2.30 respectively. Therefore, both groups of white Dutch speakers differ significantly from the Turkish-Dutch as well as from the Moroccan-Dutch

speakers. The Turkish-Dutch and the Moroccan-Dutch speakers do not differ significantly.

What happens when we include the background of the interlocutor (white Dutch, Turkish-Dutch and Moroccan-Dutch)? Figure 2 summarizes the behavior of the three groups of participants with inter-ethnic ties (D, T and M) in each of the three conditions.

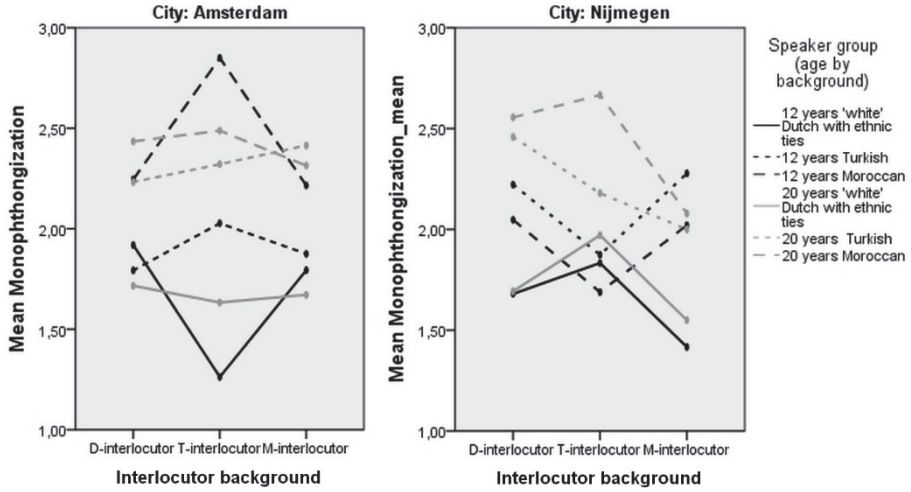


Figure 2 Mean values of monophthongization in Amsterdam (left) and Nijmegen (right) for the three interlocutor conditions: D is white Dutch, T is Turkish-Dutch, and M is Moroccan-Dutch. The values are given for six groups, i.e. age by speaker background

For all groups of Amsterdam speakers, the values for the monophthongization indexes in interactions with white Dutch interlocutors hardly differ from those in interactions with Moroccan-Dutch ones. Talking with somebody with a Turkish-Dutch background, however, triggers either the lowest monophthongization indexes (12 and 20-year-old white Dutch speakers) or, on the contrary, the highest ones (almost all other groups of speakers).

For all groups of Nijmegen speakers in conversations with Moroccan-Dutch interlocutors, the monophthongization indexes are lower than in conversations with white Dutch interlocutors – except for the 12-year-old Turkish- and Moroccan-Dutch speakers; for these groups the pattern is reversed. In Nijmegen, the patterns never go monotonically from white Dutch, via Turkish-Dutch to Moroccan-Dutch interlocutors – except for the 20-year-old Turkish-Dutch speakers. The 20-year-old Moroccan-Dutch as well as both the 20-year-old and the 12-year-old white Dutch speakers reach the highest monophthongization values while talking with Turkish-

Dutch interlocutors.

It is remarkable that in this respect the Nijmegen white Dutch speakers give the opposite picture of Amsterdam white Dutch speakers. The same also holds for the 20-year-old Turkish-Dutch speakers. It is also striking that the Amsterdam 12-year-old white Dutch participants show clear divergence when paired with the Turkish, while the 12-year-old Moroccans show clear accommodation to the Turkish. The 20-year-olds, in contrast, do not dissociate or accommodate to the same extent.

Only in the case of the 20-year-old Turkish-Dutch speakers in Amsterdam does the variation in the monophthongization of /ɛi/ in speech towards the various groups replicate the relative production of each of those groups. Clear evidence of accommodation in the sense of audience design (Bell 1984) in other groups is scarce, however. The patterns are sometimes hard to interpret because there tend to be several effects at the same time and complex interaction effects.

In the last step of the analysis, we applied mixed model analysis on the data of all speakers and all four social variables (age, city, speaker background, interlocutor background). We started the analysis by using speaker and citation forms as random variables, with random intercepts. The resulting model had an AIC of 3328.816. Including all social variables and their interactions gave an AIC of 3288.840. Removing the four-way interaction to simplify deteriorates the AIC value which becomes 3291.457. That means that this interaction needs to be included in the model plus all other interactions and main effects. The main effects of interlocutor $F(2,1297) = 1.355$, $p = .258$, and city ($F<1$) are not significant. The main effects of speaker background ($F(3,39) = 17.870$, $p=.000$) and age ($F(1,35) = 4.945$, $p=.033$) were significant (as in the ANOVA). Three complex interaction effects produce a significance value of less than .10. These are the three-way interaction between age, city and interlocutor background ($F(2,1297) = 2.612$, $p=.074$ – not significant beyond .05), the three-way interaction between city, speaker background and interlocutor background ($F(4,1298) = 4.005$, $p=.003$) and the four-way interaction between all four social variables ($F(4,1298) = 1.955$, $p=.009$). The effects confirm the impression of specific age- and city-related style shifting (adaptation to the background of the interlocutor, the background of the interlocutor returns in all three interaction effects).

Adding the four linguistic factors does not change the picture of the social variables. The AIC improves the most by including three linguistic factors (3283.153). Word class was left out. The significant effects were IJ-EI ($F(1,683) = 15.087$, $p=.000$), open vs. closed syllable ($F(1,683) = 5.730$, $p=.017$, and LOG frequency ($F(1,834) = 6.848$, $p=.009$). As to the nature of the effects: there is more monophthongization of etymological <ei> than <ij>, but this effect may well be

brought about by single <ei> items such as *eigenlijk*, 'really, actually', which generally attract monophthongization in modern spoken Dutch. As regards the syllable position: there is more monophthongization in closed position than in open position, which makes sense from a phonotactic point of view: the second element of the diphthong is a glide which can serve as a semi-consonant closing the syllable. Monophthongization in this context would result in a long lax vowel in open position, which is a highly marked structure. Finally, monophthongization increases with the frequency of usage of an item; in so far as monophthongization is reduction, this effect is in line with usage-based models such as Bybee (2001, 2006). The fact that word class, i.e. the distinction between content words and function words, does not appear to have a significant effect may be a consequence of the fact that this distinction generally coincides with a distinction in average frequency of usage - with function words being far more frequent than content words (Frisch 2011).

Statistically, it turns out to be worthwhile to include the linguistic factors, although the model improvement is not spectacular. If one looks in more detail, it is hard to interpret the results and the significant effects seem to run parallel to the intricate effects related to the background of the interlocutor. That would lead to the conclusion that specific words are more accessible to style shifting than others. The corpus we have is unfortunately too small to investigate these interesting word-bound effects in more detail.

3.5.2 Height

Figure 3 summarizes the outcomes for the two cities, Amsterdam vs. Nijmegen, split up for speaker background and age. The most outspoken social effect in Figure 3 is the effect of city. For all groups, we find that the Amsterdam speakers use more open first, prominent elements of the diphthong (i.e. have a higher mean value) than the Nijmegen ones. Dutch speakers without strong inter-ethnic ties (C) in Amsterdam have a mean value of 2.81 and the C-group in Nijmegen 2.30. The Amsterdam Dutch speakers with strong inter-ethnic ties (D) have a mean value of 2.94, the C-group in Nijmegen 2.25. For the Turkish-Dutch speakers, the values are 2.71 for Amsterdam and 2.04 for Nijmegen. The values for the Moroccan-Dutch speakers are 2.52 and 1.88 respectively. Hence, for lowering, the overall pattern is white Dutch > Turkish-Dutch > Moroccan-Dutch.

With but two exceptions (Amsterdam Turkish-Dutch and Nijmegen Moroccan-Dutch), the 20-year-old speakers have higher height values (i.e. produce lower onsets) than the 12-year-old speakers.

The ANOVA results confirm our impression. There is a strong effect for city ($F(1,35) = 126.825$, $p=.000$, partial eta squared = .784). Age ($F(1,35) = 17.575$,

p=.000, partial eta squared = .334) and speaker background (F(3,35) = 10.277, p=.000, partial eta squared = .468) are significant as well. In addition, two significant interactions were found. The first one is the interaction between age and speaker background (F(3,35) = 6.619, p =.001, partial eta squared = .362), the second one is the three-way interaction between age, city and speaker background (F(3,35) = 5.487, p=.003, partial eta squared = .320). The interaction between the age and the background of the speaker suggests more growth among the two white Dutch groups than among the Moroccan- and Turkish-Dutch: it seems that, while they grow older, there is a stronger increase in lowering of the diphthongal onset among the two groups of white Dutch speakers than among the Moroccan- and Turkish-Dutch speakers. If, however, we interpret the patterns from the point of view of apparent time change, we establish that in fact there appears to be a development away from lowering and this tendency is stronger among the white Dutch speakers than among the Moroccan- and Turkish-Dutch speakers. The three-way interaction adds the following pattern to the one revealed by the two-way interaction: in Amsterdam, it is only the Moroccan-Dutch speakers whose lowering decreases somewhat in apparent time, while in Nijmegen only the Turkish- Dutch speakers show some apparent time decrease in lowering.

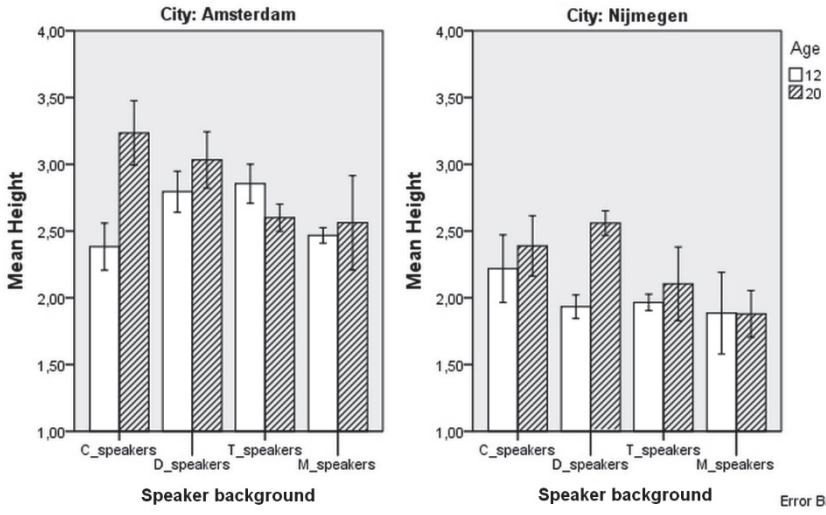


Figure 3 The mean height scores in Amsterdam (left) and Nijmegen (right), split up for background of the speaker and the two age groups; the higher the values the more lowering. The C and D speakers are the white Dutch speakers with no inter-ethnic ties (C) and strong inter-ethnic ties (D). The T speakers are Turkish-Dutch; the M speakers Moroccan-Dutch. The speaker variation within the groups is indicated by the error bars

What happens when we include the language background of the interlocutor (white

Dutch, Turkish-Dutch and Moroccan-Dutch)? Figure 4 summarizes the behavior of the three groups of participants with inter-ethnic ties (D, T and M) in each of the three conditions.

We see less obvious patterns in relation to the background of the interlocutor than for monophthongization. In both cities, the 20-year-old white Dutch have the highest scores, in particular in Nijmegen, and they perform similarly in the three conditions. That seems to also apply to the other groups.

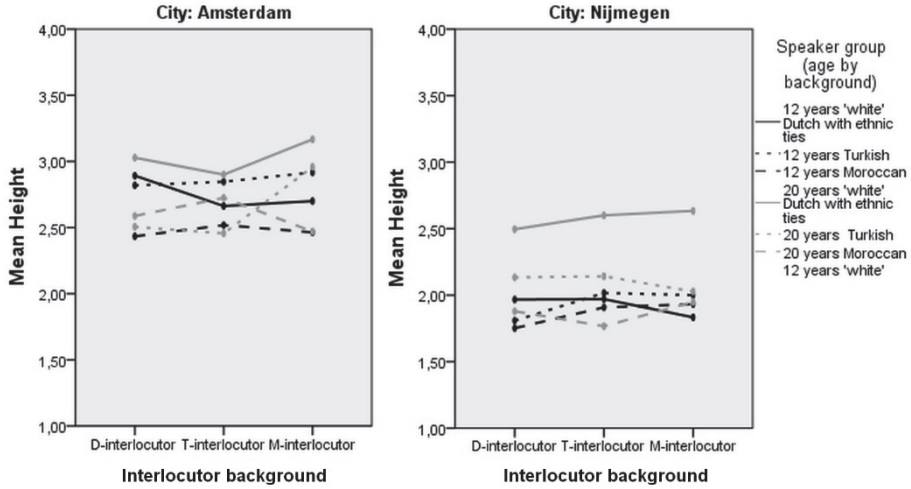


Figure 4 Mean values of height in Amsterdam (left) and Nijmegen (right) for the three interlocutor conditions: D is white Dutch, T is Turkish-Dutch, and M is Moroccan-Dutch. The values are given for six groups, i.e. age by speaker background

The results from the mixed models analysis nicely confirm this picture. The AIC value of the model with two random factors (speaker, citation form) is 3299.373. Including the four social factors and their interaction, the AIC improves to 3258.971. Deleting all effects where interlocutor background is involved improves the AIC further to 3245.199. That has the consequence that the same effects remain as we had tested in the ANOVA. The results are in fact the same. Three main effects are significant: city ($F(1,36) = 115.355$, $p=.000$), age ($F(1,36) = 15.678$, $p=.000$), and speaker background ($F(1,36) = 10.237$, $p=.000$), as are two interaction effects, i.e. age by speaker background ($F(1,36) = 5.629$, $p=.003$) and age by city by speaker background ($F(1,36) = 4.933$, $p=.006$). We have discussed these results in relation to the ANOVA outcomes.

Adding the four linguistic factors does not change the picture of the social variables. The AIC increases (3292.857), which means that the model with only the social factors included is the best one. We found no additional linguistic effects.

3.5.3 Variants

In the previous sections, we examined height and monophthongization separately. When combining the two features (four height levels and three levels of monophthongization), there are 12 possible variants of /ɛi/. In this section, we explore the occurrence of these 12 variants. We used data from the Turkish-Dutch, the Moroccan-Dutch and the Dutch speakers with strong interethnic ties as well as the Dutch with no inter-ethnic ties. At least three out of the 12 variants were used by every speaker, with a mean of 5.9 different variants. The highest number of different variants per speaker was 10 (out of 12), but this speaker used about half of them only once.

The question arises as to whether there are any group-specific effects in the distribution of the variants. A PROXSCAL multidimensional scaling (MDS) procedure with two dimensions was carried out on the percentages of used variants to find similarities between (groups of) speakers. The MDS analysis was carried out with two dimensions, as the elbow in the Scree test (to detect the number of dimensions after which the explained variance does not increase significantly) showed that two dimensions were appropriate.[8]

Figure 5 reflects the MDS analysis with each point representing one of the 51 speakers. The closer speakers are mapped, the more alike they are. Each speaker is marked for city (i.e. grey for Amsterdam, white for Nijmegen) and a white Dutch (round point) or Turkish-/Moroccan-Dutch background (square point).

The two more or less diagonal lines divide the speakers with the lowest means from the speakers with the highest means on Monophthongization and on Height.

Figure 5 features four ‘groups’: a white Dutch group from Amsterdam, a white Dutch group from Nijmegen, a Turkish-Dutch and Moroccan-Dutch group from Amsterdam, and a Turkish-Dutch and Moroccan-Dutch group from Nijmegen. In Figure 5, the Dutch are situated more on the top right-hand corner of the figure, while the Turkish-Dutch and Moroccan-Dutch speakers are more on the bottom left-hand corner. This confirms the outcomes of the mixed model analyses (discussed in Sections 3.5.1 and 3.5.2 above) as these analyses made clear that the white Dutch with strong and the white Dutch with no inter-ethnic ties form one group separate from the Turkish-Dutch and Moroccan-Dutch speakers who form another group.

The mixed model analysis for height also showed a clear main effect of city with the Amsterdam speakers using more open diphthongal variants than the Nijmegen speakers, who mainly use non-open diphthongal variants. This is reflected in Figure 5 where most Nijmegen speakers occupy the top of the figure and the Amsterdam

[8] The Stress-I was 0.13898, which is sufficient (Borg & Groenen 2005).

ones the bottom.

To find out what characterizes each of the four groups, we divided the speakers according to their mean scores on monophthongization (Section 3.5.1) and height (Section 3.5.2). This resulted in four groups: a group of speakers who can be characterized as using mainly diphthongs and open variants; a group of speakers who can be characterized as using diphthongs and non-open variants; a group of speakers who can be characterized as using mainly monophthongs and open variants; and a group of speakers who can be characterized as using monophthongs and non-open variants.

The four groups and the four areas match quite closely, as can be seen in Figure 5. This is even more obvious when we look at Table 3, in which the numbers of speakers per group are crossed with the numbers of speakers per 'area'.

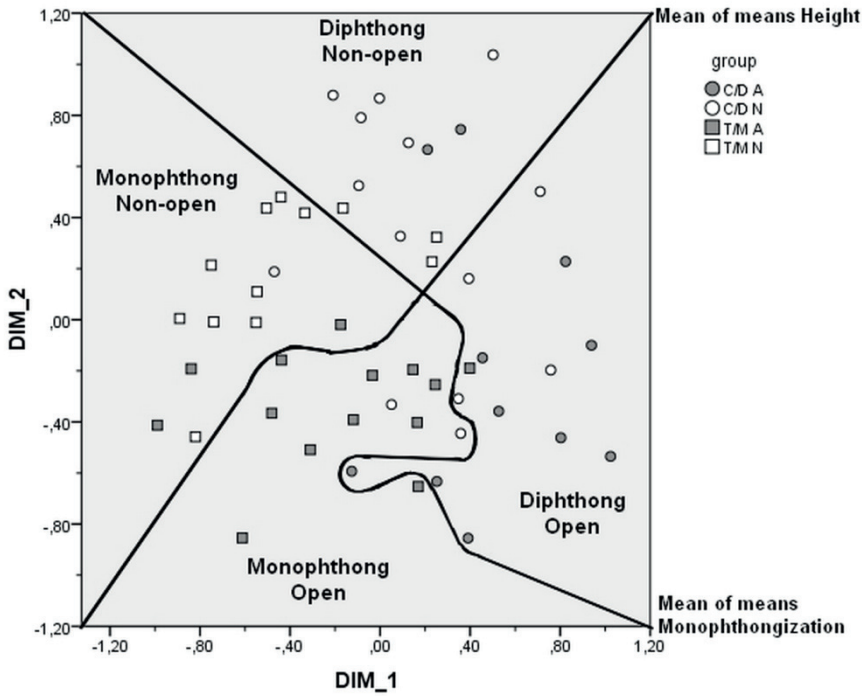


Figure 5 Positioning the speakers in a two-dimensional space based on the MDS analysis of the usage (in percentages) of variants. A = Amsterdam, N = Nijmegen, C = white Dutch without ethnic ties, D = white Dutch with ethnic ties, T = Turkish-Dutch, M = Moroccan-Dutch

Apparently, the white Dutch speakers mostly use diphthongal variants which are closer to the standard, whereas the Moroccan- and Turkish-Dutch mostly use monophthongal variants which are characteristic of the 'traditional' urban dialects.

The Amsterdamers of all backgrounds use more open variants and the Nijmegen speakers of all backgrounds use more non-open variants which is in line with the 'traditional' urban dialects.

Table 3 Comparison of the four groups of speakers and the four areas of Figure 5. A = Amsterdam, N = Nijmegen, C = white Dutch without ethnic ties, D = white Dutch with ethnic ties, T = Turkish-Dutch, M = Moroccan-Dutch

	diphthong + open	diphthong + non-open	monophthong + open	monophthong + non-open	Total
C/D A	**9**	2	0	0	11
C/D N	4	**7**	2	1	14
T/M A	1	0	**10**	3	14
T/M N	0	3	0	**9**	12
Total	14	12	12	13	51

3.6 Conclusion and discussion

The variation in the realization of the diphthong /ɛi/ has several dimensions; monophthongization is sensitive to etymological distinction between <ei> than <ij> (which is phonological in several groups of Dutch dialects), the position vis-à-vis the right syllable boundary and the frequency of occurrence of the citation forms. The lowering of the first, prominent element of the diphthong does not appear to be affected by linguistic variables at all. Linguistically, this finding makes sense in that monophthongization basically changes the phoneme into a sound which is homophonous with the monophthongal phonemic counterparts /ɛ/ and /a/, whereas lowering merely has phonetic consequences.

There are also complex style effects in the variation in the realization of the diphthong /ɛi/ which are connected with the background of the interlocutor. For the monophthongization of /ɛi/ this style-as-accommodation effect is found only in the speech of the 20-year-old Turkish-Dutch speakers from Amsterdam. For lowering, only the 20-year-old Turkish-Dutch cohort from Nijmegen show this style effect.

In Section 3.1, three hypotheses were developed concerning the relative use of the dialectal, the standard and the younger substandard variants of the diphthong by the members of the various groups in both cities. The outcomes presented in Section 3.5.1 (monophthongization), 3.5.2 (height) and 3.5.3 (the overall picture regarding the distribution of the main variants) above allow these hypotheses to be tested.

According to hypothesis 1, the white Dutch boys will target the more prestigious standard forms of the standard Dutch (local divergence, upward convergence toward the standard norm). This hypothesis is borne out for monophthongization (Section 3.5.1), but not for height (Section 3.5.2). Underlying this development may be either generally ongoing processes of dialect leveling or a desire for upward mobility on the

part of the speakers – or both.

Hypothesis 2 says that the traditional dialect variants are adopted by the ethnic groups (local convergence towards the socially 'low' urban accent). Urban accent variants, thus, become part of the ethnolects. This hypothesis, too, is supported as far as the monophthongization is concerned (Section 3.5.1).

The outcomes of the MDS, presented in Section 3.5.3, show that there is indeed a very clear tendency towards a double resetting of the social distribution of the variation in the realization of /ɛi/: local dialect variants appear to be changing from sociolectal into ethnic markers.

Hypothesis 3 claims that there will be no differences between both ethnic groups, resulting into a multi-ethnolect (inter-ethnic convergence). This hypothesis is corroborated in all relevant respects; the Moroccan- and Turkish- Dutch speakers are united in their embracing the older local dialect variants of the diphthong.

In sum, there is a social redistribution going on of urban dialect features, which had developed into a sociolect feature over the past generations. The dynamics revealed on the basis of our data show that the 'traditional' urban dialects are being recycled to become an ethnolect marker. Scholars such as Rampton (2011) and Jaspers (2011) have paid attention to the mechanism that linguistic features with urban and socially low indexicalities are taken up by speakers with migration backgrounds.

In Section 3.2 above, we discussed various definitions of the ethnolect concept; in the same section, we also briefly sketched different approaches to ethnolectal variation. Our language-centered, quantitative approach has revealed ongoing changes in the overall social distribution of variation in the realization of the /ɛi/ diphthong, changes in which the endogenous urban dialects and the ethnolects are involved in a game of musical chairs - showing how intricately dialect variation and ethnolectal variation can be entwined.

In Van Meel, Hinskens & van Hout (2013), we presented our findings for variation in the realization of /z/ by the same speakers. Unlike /z/, and especially the 'exotic' dental variants in its realization, /ɛi/ seems not to be a main candidate sound to become an ethnolect marker, being first of all a dialect marker and a marker of ongoing change in standard Dutch. Nevertheless, it has become part of a process of redistributing the socially emblematic value of different variants, in which immigrant youngsters play the role of the savers of the local urban dialects.

Though statistically significant, both the linguistic and the stylistic effects found are weak and somewhat inconsistent and this may be a consequence of the modest number of speakers studied. A desideratum for further research is, therefore, to

broaden the empirical basis of this study by including more speakers for each of the relevant groups defined by the extra-linguistic variables city, cultural background and age group.

4 Variation in a tense/lax vowel pair in Dutch youngsters with different ethnic backgrounds

Edited from: van Meel, Linda, Frans Hinskens & Roeland van Hout. submitted. Variation in a tense/lax vowel pair in Dutch youngsters with different ethnic backgrounds.

Abstract

How do young bilingual speakers of current Turkish and Moroccan ethnolects of Dutch deal with phoneme contrasts that do not exist in their heritage languages and that are at the same time subject to regional and social variation in the Dutch speech community at large, such as that between Dutch phonemes /a:/ and /ɑ/? Data from speakers from the Amsterdam and Nijmegen urban areas were analyzed and compared. Two variable properties of /a:/ and /ɑ/ were examined: (1) duration, and (2) place of articulation.

We found clear differences between the two urban areas (regional effect) and between the Amsterdam groups (social effect). In addition, we found variation dependent on the interlocutor (socio-stylistic effect) and the linguistic context. The main source of social and linguistic variation was place of articulation, length remaining the primary distinctive feature of the vowel pair. Thus, the expected heritage language effect did not manifest itself.

4.1 Introduction

Ethnolects, language varieties which originate among members of specific ethnic groups, are a relatively new domain in the study of language contact and bilingualism. The present contribution focuses on the analysis of linguistic variation from a language centered, sociolinguistic perspective, using quantitative tools to analyze patterns of linguistic variation. The analyses concern variation in the realization of the Dutch /a:/-/ɑ/ vowel pair contrast by young speakers from the urban areas of Amsterdam and Nijmegen, representing second generation Moroccan-Dutch and Turkish-Dutch as well as endogenous 'white' Dutch. This vowel pair is hard to learn for adult second language learners. El-Aissati, Boumans, Cornips, Dorleijn & Nortier (2005:154) observed that two first generation Moroccan woman did not produce any distinction in their L2 Dutch. They also studied a first generation Turkish woman who produced all tensed vowels as their lax counterparts, preserving their distinction by making them long. This woman pronounced *naartoe* 'towards' (standard Dutch [na:rtu]) as [nɑ:rtu] where the [ɑ] has a longer duration having the length of a tensed vowel (El-Aissati et al. 2005:156). Van Krieken (2004) examined three second generation Moroccan participants. She noticed that two participants shortened long vowels overall, including /a:/.

The central question of this article is how young bilingual speakers of Turkish and Moroccan origin deal in their Dutch with the phoneme contrast /a:/ and /ɑ/ that does not exist in their heritage languages and that is at the same time subject to

regional and social variation in the Dutch speech community at large. The data are extracted from recordings of natural conversational speech in both in-group and out-group contact situations. The findings serve to address two general questions. The first general question we want to address is to what extent patterns of variation in /a:/ and /ɑ/ are rooted in interference from the original language of the ethnic group in question ('substrate effects'), in properties resulting from processes of second language acquisition or in endogenous non-standard varieties (typically urban dialects). The second general question addresses the place of the vowel pair in the verbal repertoires of its speakers, including the issue regarding to what extent speakers of an ethnolect can shift towards more standard varieties and towards (non-ethnic) non-standard varieties.

Ethnolects have hardly been systematically studied in linguistic detail as yet, neither for the Dutch situation nor for most other language areas. Before we present our empirical study, the ethnolect concept as well as some of the most pervasive topics in its study will be briefly discussed in the following subsections, including the context of the present study.

4.1.1 Defining ethnolect

As the essence of ethnolects is still largely unknown, most definitions of the concept of ethnolect are stipulative in the sense that they are conventions on what is meant by that notion. Clyne (2000:86), for instance, has defined ethnolects as ‘varieties of a language that mark speakers as members of ethnic groups who originally used another language or distinctive variety’. Androutsopoulos (2001:2) defined an ethnolect as ‘a variety of the majority language (or ‘host language’) which is used and regarded as a vernacular for speakers of a particular ethnic descent and is marked by certain contact phenomena’. According to Auer (2003:256), ‘an ethnolect is a way of speaking (a style), which by the speakers themselves or by others is associated with one or more non-German ethnic groups’ [our translation]. In Auer’s conception, an ethnolect (also) concerns grammar, as opposed to the known innovations of youth language. Muysken (2013) describes more such oppositions. In his view, ethnolects are more or less stable and their usage is at most semi-conscious. Ethnicity plays an inherent role, and the features involved are in the domains of phonology and syntax. Youth language, on the other hand, is dynamic and lexical, and its usage is conscious. It deviates from other varieties mainly lexically; other components (phonetics, phonology, morphology, syntax) hardly play an essential role in setting it apart from ethnolects.

Features which originated in the language contact situation underlying the development of a specific ethnic variety sometimes spread to other ethnic groups to

become (what has been referred to as) 'multi-ethnolect' features (Clyne 2000; Quist 2008; Wiese 2009, 2013). Multi-ethnolect features can also be stabilized second language acquisition phenomena and they, thus, need not be specific to any ethnic group and their heritage L1. An example is variation in the marking of Dutch grammatical gender (Hinskens et al. in preparation).

In connection with the functional dimension, the question arises as to whether ethnolects are *Mediums for Inter-ethnic Communication* or rather *Mediums for Community Solidarity*, in Baker's (2000) terminology. In the latter case, ethnolects will probably function mainly or merely as in-group codes; in that case, the emblematic value of the ethnolectal variants, which are often quite distinct from the prestigious norm, is mainly defined by their signaling ethnic identification and solidarity. This is line with Benor's (2010:160) concept of ethnolinguistic repertoire 'as a fluid set of linguistic resources that members of an ethnic group may use variably as they index their ethnic identities'.

In general, two complementary approaches to the study of ethnolectal variation can be distinguished: the language structure-centered and the ethnographic approach. Whereas the ethnographic approach conceives language systems as infinite resources from which speakers may freely[1] choose to construct their identity, the language structure-centered approach tries to disentangle the laws, generalizations and restrictions on these resources, with a distinct emphasis on linguistic embedding. The language structure-centered approach is marked by terminology such as 'ethnolect', 'multi-ethnolect' and 'multicultural variety'. The approach, which highlights features of linguistic structure, their origin and distribution, is quantitative, often in the Labovian tradition. The features' patterns of use are usually viewed from a rather macro-social perspective (e.g. Cheshire et al. 2011; Hoffman & Walker 2010).

4.1.2 Dutch ethnolects

A number of different Dutch ethnolects can be distinguished beginning in the nineteenth century and up until the last quarter of the twentieth century.[2] Some of these have been documented relatively well, others much less so. Two examples of historical ethnolects are Jewish Dutch and Indonesian Dutch.

As far as the recent past is concerned, there is a steadily growing Chinese community in the Netherlands ever since the second quarter of the 20th century, their

[1] I.e. without discernable internal conditioning – as in the structuralist concept of 'free variation'.
[2] The past and present-day emergence of Dutch ethnolects is detailed in Hinskens (2011) and Muysken (2010; 2013). Cf. Den Besten & Hinskens (2005: 289-293) on Jewish Dutch and Hinskens & Muysken (2007) on (among other things) Jewish Dutch and the Chinese and Surinamese communities in present-day Amsterdam.

main heritage languages being Cantonese and Hakka. In 1988 the group was estimated at a total of around 45,000 people (van de Berg & Pieke 1991); recent statistics mention some 80,000 individuals, spread out over towns all over the country without any particular geographical concentration. Very little scholarly attention has been paid to their Dutch.

After World War II considerable numbers of Ambonese people from Indonesia settled in the Netherlands. They used to speak Melaju Sini (Dutch name: 'Ambonees'), a Malay based creole language which is mainly spoken on the Moluccan islands Ambon and nearby Ceram (which have been a part of Indonesia since 1949), as well as Moluccan Malay Dutch (1920–). This ethnolect has been reported to be dwindling (Tahitu 1989).

Surinamese Dutch (1900–) is one of the modern ethnolects. After approximately 1975 immigration from Surinam has gradually reached significant proportions. The majority of the speakers of Antillean Dutch (1950–) come from the Caribean island of Curaçao. More generally, with respect to the immigration from the former Dutch colonies, the Surinamese and Antilleans are not only the most recent (and still ongoing, especially as far as the Antilles is concerned), but also among the most influential cultural groups. In contrast to Jewish Dutch and Indonesian Dutch, which according to De Vries (2005) are gradually fading away, Curaçao and Suriname ethnolects of Dutch are alive and well. These ethnolects are spoken by people who themselves and/or whose ancestors came from present and former Dutch colonies.

In this contribution, we present findings based on data from speakers of Moroccan and Turkish ethnolectal varieties of Dutch living in the cities of Amsterdam and Nijmegen. These ethnolectal varieties of Dutch originated in labor migration which gained momentum in the 1970s. Both ethnolectal varieties have been investigated by several researchers (e.g. El-Aissat et al. 2005, van Krieken 2005). Some researchers have investigated these new ethnolects in relation to the notions of stylization and identity and from a rather ethnographic perspective (cf. Flanders: Jaspers 2008; Netherlands: Nortier & Dorleijn 2008).

4.1.3 The urban speech communities

The dialects spoken in Amsterdam and Nijmegen have become urban sociolects as a result of the fact that by the end of the 19th century standard Dutch started to take root in oral usage in the higher status groups. Nowadays, the Amsterdam and Nijmegen urban dialects are spoken in their most pronounced form in the low-income neighborhoods (cf. for Amsterdam: Brouwer 1989; Schatz 1986; for Nijmegen: van Hout 1989, 1999) and their prestige is low. The urban linguistic situations are marked by socially stratified linguistic continua between the urban dialect and standard

Dutch.

Former immigrants from Turkey and Morocco and their families typically live in densely populated neighborhoods with cheap housing, low incomes, high unemployment rates and reduced access to infrastructure. Growing up in these areas situated near the bottom rung of the socio-economic ladder brings the second and third generation immigrants into contact with peers of Dutch descent who use urban dialect as their native speech in the neighborhood as well as in school. This situation of long-lasting, intensive exposure seems to be an ideal context for youngsters with a different ethnic background to acquire the surrounding local urban dialect. And judging from the variation in their Dutch speech, this is in fact what they do – as we will explain in the next subsection.

At present, 51% of the population of Amsterdam and 25% of Nijmegen are *allochtonen*, people of whom at least one parent was born and raised abroad.

4.1.4 Earlier findings from the Roots of Ethnolects project

The study we present here is part of a larger research project 'The roots of ethnolects. An experimental comparative study'. The project concentrates on the emergence, position and social spread of young ethnolects of Dutch in Amsterdam and Nijmegen. Earlier findings were presented in van Meel, Hinsken & van Hout (2013, 2014) and Hinskens (2011).

Van Meel, Hinskens & van Hout (2014) focus on the Dutch diphthong /ɛi/ that can be pronounced in varying ways. A first relevant distinction is the one between the urban dialects of Amsterdam and Nijmegen. In Amsterdam, this diphthong is subject to monophthongization and lowering, leading to realizations as [æ:] or [a:]. In Nijmegen, the diphthong is only subject to monophthongization, leading to the variant [ɛ:]; both variants have low overt (and no covert) prestige. A second relevant distinction concerns the spoken standard variety. From the early 90's onwards, a new variant of the diphthong /ɛi/ has been observed in colloquial standard Dutch: the lowered, diphthongal variant [ai]. All this means that ethnolect speakers have a pool of variants they can 'choose' from: the traditional standard Dutch variant [ɛi], the new, expanding variant [ai], and the local monophthong variants marking the dialects of Amsterdam and Nijmegen respectively. The outcomes from our quantitative analyses show that 'white' Dutch speakers appear to favor the new substandard or even standard realizations, while Turkish- and Moroccan-Dutch speakers are taking over the older urban dialect variants which had developed into markers. The traditional dialect variants of this diphthong thus undergo a complex, long-term sociolinguistic evolution: about a century ago they developed from dialect into sociolect features, now they seem to be evolving from sociolect into ethnolect markers.

Hinskens (2011) and Van Meel, Hinskens & van Hout (2013) deal with the ethnolect speakers' variation in the realization of the voiced sibilant /z/. The results show that /z/ in word-initial position has several traces of interference from the original languages of one of the ethnic groups under study. The dental, voiced realization of /z/, which occurs variably in the Dutch ethnolects of Turkish-Dutch and Moroccan-Dutch speakers in the cities of Nijmegen and Amsterdam, has its origin in the Moroccan languages and is not part of the dialectological and/or sociolinguistic patterns of variation of traditional endogenous Dutch. This dentalized /z/ has noticeably and demonstrably more frication and it therefore sounds 'sharper' than the traditional Dutch variant, which is alveolar. Apart from the place of articulation, we also investigated voicing, as in a large part of the language area (including the Amsterdam and Nijmegen regions) /z/ is involved in processes of devoicing in standard Dutch. We established that the Turkish and Moroccan ethnolectal varieties of Dutch are marked by significantly less devoicing. Moreover, the speakers of these ethnolects have reshaped the phonological conditioning of the voiced /z/ to include environments following an obstruent – where no endogenous variety of Dutch would ever show any voicing.

4.1.5 The research questions

In the present study, we use the same corpus as in our earlier studies to investigate how speakers of current Dutch ethnolects deal with Dutch segmental contrasts which do not occur in the heritage languages involved, and which are marked by intricate regional and social stratification patterns in varieties of Dutch. A contrast that unequivocally meets these criteria is that between Dutch long and short (phonologically analyzed as tense versus lax) vowels, such as /a:/ - /ɑ/ as in e.g. *maan* 'moon' – *man* 'man', *maat* 'measure; buddy' – *mat* 'mat', and the like. Such a contrast does not occur in Turkish, Moroccan Arabic and Berber (see section 2). Therefore, its absence from the heritage languages involved is a good basis to establish any second language acquisition effects.

While the variation in the realization of /z/ is rooted in substrate effects and the variation in the realization of /ɛi/ originates in the surrounding urban dialects / sociolects, in studying the variation in the realization of the /a:/-/ɑ/ contrast we hope to be able to test for second language acquisition effects. After all, unlike Dutch, neither Turkish nor Moroccan Arabic nor Berber have length contrast in their vowel inventories (see again section 2). And in connection with this, it must also be established that (in contrast to Dutch, which is a stress-timed language) Turkish, Moroccan Arabic and Berber are syllable-timed systems (Abercrombie 1967; Pike 1945).

The speech of some Turkish- and Moroccan-Dutch speakers in our sample has a “staccato” ring to it and this could well be connected with this typological distinction which defines ‘a complex of prosodic-phonetic features’ (Keim 2007), including phonetic reduction processes. It may well be the case that the salient rhythm is brought about in part by a tendential isochrony of syllable length – in particular vowel length. Vowel duration will be part of the focus of our analyses of /a:/ and /ɑ/ and on the basis of the data we hope to address the two research questions mentioned in the beginning.

The first general question we want to address is related to the embedding of the patterns of variation in /a:/ and /a/ in relation to the original languages, second language acquisition and the urban dialects. More specifically we pose the following specific research questions:

- Are there any systematic differences between the Turkish- and Moroccan-Dutch on the one hand and the ‘white’ Dutch speakers on the other in the way they realize vocalic tense-lax contrasts between /a:/ and /ɑ/? If so, then do the differences concern vowel quantity (length, duration) or quality (place of articulation) or both?
- Is the internal conditioning (in terms of segmental environment) of the differences in duration and place of articulation identical for the ‘white’ Dutch on the one hand and the Turkish- and Moroccan-Dutch on the other?
- What is the social distribution of the variation in the realization of both contrasting vowels?

The second general question addresses the place of the vowel pair in the verbal repertoires of its speakers. The specific research question we want to answer is:

- Do the patterns of variation found play a role in the style continuum in the verbal repertoires of the speakers?

The design of our database (which will be presented in Sections 4.1 and 4.2 below) enables us to address these various issues empirically, by studying the distribution of the variation in /a:/ and /ɑ/ in extralinguistic dimensions while taking into account the role of linguistic conditions, and by considering the way variation plays an active role in face-to-face interaction (style shifting dependent on the background of the interlocutor).

In section 2, we first will discuss the different realizations of the /a:/-/ɑ/ contrast in standard Dutch, of the relevant variants in the Amsterdam and Nijmegen dialects, and their nearest neighbors in the heritage languages at issue. The methods are discussed in Section 3, followed by a presentation of the results in Section 4. Section 5 contains the discussion and conclusions.

4.2 /a:/ and /ɑ/ in the languages involved

The information about /a:/- and /ɑ/-like segments in the language varieties involved is summarized in Table 1. The details are discussed in the following subsections. It turns out that none of varieties involved has a phonemic distinction in the lowest part of the vowel triangle except for Dutch and its urban varieties. Nevertheless, there is a lot of (allophonic) length and place variation in the low vowels in all varieties. In the two Dutch urban dialects involved, many variants are found for place of articulation, but not for length. Length seems to be the more solid feature for Dutch and its varieties, apart from general processes of vowel lengthening and vowel reduction in general (cf. Booij 1995).

Table 1 Overview of relevant phonemes and their features in the language varieties involved; PoA = place of articulation

Language	Phonemes, features	Variability
Standard Dutch	/a:/ and /ɑ/ tense vs. lax, different in PoA and length	
Amsterdam	/a:/ and /ɑ/ tense vs. lax, different in PoA and length	• PoA variants for both phonemes • stable length distinction
Nijmegen	/a:/ and /ɑ/ tense vs. lax, different in PoA and length	• PoA variants for /a:/ • stable length distinction
Turkish	/a/, also represented as /ɑ/	• allophonic variation in PoA and length • long variant in loan words
Moroccan-Arabic	/a/	• allophonic variation in PoA and length
Tarifit (Berber)	/a/	• allophonic variation in PoA • long realization in closed syllables

Morocco is a multilingual country with Moroccan Arabic and Berber languages used as mother tongue(s) by the Moroccans (El-Aissati & E-rramdani 2001; cf. El-Aissati et al. 2005). Moroccan Arabic consists of several dialects which are mutually intelligible. The Berber languages can be divided into three main language groups that are not mutually intelligible: 1. Tarifit (also called Riffian or Rif Berber), 2. Tachelhit (also written as Tachelhiyt or Tashelhiyt), 3. Tamazight (Central Atlas). El-Aissati et al. (2005:150) estimated that some 60 percent of the Moroccans in the Netherlands speak Tarifit, 10 percent to speak Tachelhit and the remaining 30 percent to speak a Moroccan Arabic dialect as their mother tongues. In the last two subsections below, the phoneme /a/ is being discussed for Moroccan Arabic and for the Berber languages concentrating on Tarifit.

4.2.1 Standard Dutch

Standard Dutch, as spoken in the Netherlands, has nine full monophthongs, of which

/a:/ and /ɑ/ are the lowest two in place of articulation (Gussenhoven 1999). The two vowels differ in several respects. First, /a:/ is phonologically a tense and /ɑ/ a lax vowel. In Dutch, tense vowels are phonologically long (except the high ones), which also means that they occur in other environments than their counterparts, the lax vowels, which are all phonologically short. Lax vowels, such as /ɑ/, have to be followed by a tautosyllabic consonant (e.g., */pɑ/, but /pɑs/), whereas tense vowels may end a Dutch syllable (e.g., both /pa:/ and /pa:s/ are allowed). Lax vowels may also be followed by more consonants within the same syllable than tense vowels: /a:/ may be followed by only one consonant, plus /s/ and/or /t/, whereas /ɑ/ may be followed by two consonants, plus 'extrasyllabic' /s/ and/or /t/ (Booij 1995). So, /kɑlm/ is a possible word in Dutch, but /ka:lm/ is not.

It follows that in a considerable number of cases the two low vowels occur in the same phonological environment. For instance, *lat* /lɑt/ 'pole' and *laat* /la:t/ 'late' are two different Dutch words. The risk of confusing the two words seems high, because the vowels are both phonologically low and unrounded. However, Booij (1995:5) and Gussenhoven (1999) point out that /a:/ is (phonetically) more central, i.e. front, and, as mentioned above, it is longer. The acoustic study of Van der Harst, van de Velde & van Hout (2014) additionally shows that /a:/ is lower and more fronted than /ɑ/.

Van der Harst et al. (2014) only found some minor regional differences between standard language speakers in the Netherlands: /ɑ/ is slightly more front among standard speakers in the northeast of the Netherlands than among standard speakers in the southeast of the Netherlands. For /a:/, Van der Harst et al. (2014) do not report regional differences. The vowel is generally assumed not to be subject to any change in Standard Dutch (cf. van Heuven, Edelman & van Bezooijen 2002).

4.2.2 The urban dialect of Amsterdam

As described in the previous section, /a:/ in Standard Dutch is more central, i.e. front, than /ɑ/. In the Amsterdam dialect, this opposition is reversed, with /ɑ/ being more fronted than /a:/ (cf. Schouten, Crielaard & van Dijk 1998). In addition, the height difference between the two vowels, measured as the difference in the second formant, is substantially smaller than in Standard Dutch (Schouten et al. 1998:113). Schatz (1986:62, 63) observed a fronting and raising of [ɑ] to [ɛ] in Amsterdam vernacular, i.e. Standard Dutch *pan* 'pan' /pɑn/ being pronounced as [pɛn]. This seems to match the fronting found by Schouten et al. (1998). At the same time /a:/ is subject to backing, together 'with a certain degree of rounding and diphthongization' (Schatz 1986:63), i.e. *kaas* 'cheese' /ka:s/ being pronounced as [ka:ɔs]. Brouwer (1989:29) characterizes the Amsterdam vernacular /a:/ as [+low], [+back], [+round]. She distinguished four variants that range from the standard variant to the (most)

non-standard variant: [a:], [ɑ:], [ɑ.ᵊ], and [ɔ:] (Brouwer 1989:30, 32), going from front to back, with rounding.

4.2.3 The urban dialect of Nijmegen

The Dutch phoneme /a:/ has different origins in the West-Germanic vowel system. These origins were still present in the older stages of the Nijmegen dialect, but they disappeared more recently under the influence of standard Dutch (van Hout 1989). /a:/ has three variants in the spontaneous speech of Nijmegen speakers. There is the raised variant [æ:], but also the back variant [ɑ:] (the most frequent variant in the dialect of Nijmegen nowadays) and sometimes the back rounded variant [ɔ:]. /ɑ/ has no dialect-related variants, although it can be pronounced a bit more rounded than in Standard Dutch. Finally, the Nijmegen dialect is marked by stronger processes of vowel shortening and reduction than Standard Dutch (van Hout 1989:206).

4.2.4 Turkish

Turkish has a vowel /a/ which is traditionally described as low or non-high, back and unrounded (e.g. Comrie 1997), is often represented as /ɑ/ (e.g. Kiliç & Öğüt 2004; Ladefoged 2005; van Heuven & van Houten 1985). The vowel has two allophones. The first allophone is described by Lewis (2000[1967]:13) as 'a back, open, unrounded vowel, like the *a* of French *avoir* or northern English *man'*. Swift (1963:8) describes this allophone as 'a low central unrounded vowel approximating the vowel of American English *hot* '. According to Göksel & Kerslake (2005:xxii) this allophone is 'pronounced as *u* in *cup*'. The second allophone of /a/ is a more fronted one. According to Lewis (2000[1967]:14) this 'front sound of *a*' is 'verging on that of *e* (i.e. like French *être*), which can be heard in careful speakers' pronunciation of some Arabic borrowings and in the Istanbul word *anne* 'mother' (elsewhere *ana*)'. Göksel & Kerslake (2005:10) restrict the context further: 'Its fronted allophone [ạ] occurs with the palatal consonants /c/, /ɟ/ and /l/ in loan words.' The native Turkish vowel /a/ is normally short phonemically, but may be long depending on syllable position and following context (Göksel & Kerslake 2005; Kornfilt 1997; Lewis 2000[1967]). In contrast to native Turkish words, borrowed words (from Persian and Arabic) can have a phonemic long vowel /a:/.

4.2.5 Moroccan Arabic

Moroccan Arabic has a vowel /a/ that occurs in all positions within words, in combination with all consonants, and is never affected by elision and inversion. It has several different pronunciations though. Moroccan Arabic /a/ is generally fronted to phonetic [æ] in non-backing consonantal environments, i.e. when it is not next to an

emphatic consonant or one of the uvular and pharyngeal consonants *q*, *x*, *ġ*, ḥ , and ʕ (Harrell 1962; Heath 2002) and when it is in non-final position. This allophone is characterized as 'High Low' and 'Front' by Abdel-Massih (1973). /a/ then sounds like [æ] in English 'mad'.

When /a/ is preceded or followed by an emphatic consonant, it is pronounced approximately 'like the 'a' of English 'father'' (Harrell 1962:12) or even 'a little further back than the 'a' of English 'father'' according to Abdel-Massih (1973:24). This allophone is characterized as 'Low' and 'Back' by Abdel-Massih (1973:23).

Harrell (1962) distinguishes a third allophone. Before or after the uvular and pharyngeal consonants as well as at the end of words (the latter irrespective of the preceding consonant) the 'a' has a pronunciation 'intermediate between the 'a' of *mad* and the 'a' of 'father'' (Harrell 1962:12).

4.2.6 Berber

Just as Moroccan Arabic, most Berber dialects have three vocalic phonemes: /i/, /u/ and /a/ (Kossmann & Stroomer 1997). Kossmann & Stroomer (1997) state that the phonetic realization of the vowel /a/ ranges from [ɑ] to [æ]. According to Lafkioui (2007:17), the phoneme /a/ is a mid vowel that is realized as [æ] or [ɛ]. When the vowel is in contact with back phonemes, it is pronounced "more front and more open", and becomes [a] (Lafkioui 2007:17 [our translation]). The vowel positions more backwards when in contact with pharyngeal phonemes and is pronounced as [ʌ] or [ʌˑ] in closed syllables at word end (Lafkioui 2007:17). The vowel /a/ becomes longer in closed syllable i.e. [æˑ] or [ɛˑ]. The lengthening in closed syllables at word end is a characteristic of the Riffian vowel system.

4.3 Methods

4.3.1 Participants

For the current research we used the same sample of 51 youngsters as in previous studies (van Meel, Hinskens & van Hout 2013, 2014). The participants represent three different groups: Moroccan-Dutch, Turkish-Dutch and 'white' Dutch. The groups were controlled for residence (Amsterdam vs. Nijmegen), for age (10-12 versus 18-20 years old) and, in the case of the 'white' Dutch speakers, the presence (strong inter-ethnic ties) or absence (no inter-ethnic ties) of regular contacts with Turkish-Dutch and Moroccan-Dutch peers. Table 2 gives the speaker design.

All participants were born in the Netherlands and grew up in their place of residence (either Amsterdam or Nijmegen), Dutch being (one of) their mother tongue(s). The participants with a Moroccan or Turkish background have at least one parent who immigrated from Morocco or Turkey respectively.

Table 2 Overview of the research design and the number of participants

Background	Moroccan-Dutch		Turkish-Dutch		'white' Dutch			
Inter-ethnic ties?	Yes		Yes		Yes		No	
years of age	10-12	18-20	10-12	18-20	10-12	18-20	10-12	18-20
Amsterdam	3	4	3	4	2	3	3	3
Nijmegen	3	3	3	3	3	3	4	4

Due to unanticipated complications during the fieldwork sessions, there is only scant information available about the language skills and linguistic profiles of the speakers in our sample. The notion linguistic profile refers to such issues as the domains of language use, network-specific linguistic practices and the like. For 14 of the 51 speakers in the sample studied, we have no information whatsoever about their linguistic profiles; 7 of them being 'white' Dutch, 4 Turkish-Dutch and 3 Moroccan-Dutch. All Turkish-Dutch speakers reported knowledge of Turkish and all Moroccan-Dutch speakers reported to know either one or more of the Moroccan languages as shown in Table 3.

Table 3 Self-reported language skills; for all subsets of speakers, the languages are ordered identically, starting with Dutch, i.e. they do not necessarily reflect the order given by the speaker(s); (English, German, French and Spanish are foreign languages learned in secondary education)

Background	Years of Age	*n* of speakers	Language skills
Turkish-Dutch	10-12	6	Dutch, Turkish
	18-20	2	Dutch, Turkish
		4	Dutch, Turkish, English
		1	Dutch, Turkish, English, German, Spanish
Moroccan-Dutch	10-12	2	Dutch, 'Moroccan'
		1	Dutch, Moroccan Arabic
		2	Dutch, Berber
		1	Dutch, Berber, 'a little Moroccan Arabic'
	18-20	1	Dutch, 'Arabic'
		1	Dutch, Moroccan Arabic, English
		1	Dutch, Moroccan Arabic, English, German
		1	Dutch, Moroccan Arabic, English, French, Spanish
		2	Dutch, Berber, English
		1	Dutch, Berber

Reading Table 3, several things need to be kept in mind. First, 'language skills' was usually handled by the informants as a yes/no variable and there is no reliable information regarding the speakers' relative proficiencies in the languages mentioned – nor in the varieties of the languages at issue. Second, certain languages were sometimes not mentioned precisely (e.g. 'Moroccan' may refer to Moroccan Arabic, to Berber, or even to both). The fact that precise information was not always asked

for by the fieldworkers further complicates the picture. For all these reasons, we refrained from using this information in our analyses.

4.3.2 Material and data

Each participant took part in three to four rather free conversations. The speakers with a 'white' Dutch background with inter-ethnic ties as well as those with a Moroccan-Dutch or Turkish-Dutch background were recorded in at least one in-group conversation and two different out-group conversations. In an in-group situation the conversation partner was a peer with the same background (i.e. a Turkish-Dutch speaking with a Turkish-Dutch), while in the out-group situations they talked with a peer from the other two ethnic groups. The 'white' Dutch participants with no (or at best very weak) inter-ethnic ties (C group) were only recorded in in-group conversations with fellow Dutch that equally have no inter-ethnic ties. They serve as the control group for the 'white' Dutch participants with strong inter-ethnic ties (D group). Each conversation involved two peers of the same age group. The majority of the speakers attended the same school as their conversation partners, and many of them were classmates (especially so in the case of the 10-12 year olds). Preferably, the interviewer was only present at the beginning and the end of the conversation to ensure a more natural conversation. However, especially so in the case of the 10-12 year olds, additional guidance was needed to keep them talking for one hour (i.e. suggesting topics to talk about, introducing card games).

The conversations were recorded on a Marantz Professional CD recorder CDR300. Ten to 15 minutes of each recording was transcribed orthographically using the multimedia annotator Elan (cf. Brugman and Russel 2004). These transcriptions were checked by a second transcriber.

From each conversation, up to 20 /a:/-sounds and up to 20 /ɑ/-sounds in closed syllables were selected per speaker according to the divisions in Table 4, starting with the checked part of the transcriptions. If there were not sufficient words per cell (i.e. 5, see Table 4) in the checked part, the unchecked transcriptions were scanned.

Table 4 Number of targeted /a:/ and /ɑ/ sounds in closed syllable per speaker per conversation

	__[non-liquid]		__[liquid] (i.e. /l/ or /r/)	
	primary stress	secondary stress	primary stress	secondary stress
/a:/	up to 5	up to 5	up to 5	up to 5
/ɑ/	up to 5	up to 5	up to 5	up to 5

The first three minutes of each conversation were skipped to give the participants some time to get used to being taped and pay less attention to their way of speaking. We defined a set of criteria to select proper /a:/ and /ɑ/ words. The words had to

meet the following criteria:

1) /a:/ or /ɑ/ must have either primary or secondary stress.
2) To ensure the data were not biased by specific high frequency words, a word was selected at most two times for a given speaker in a given conversation. However, a word was only selected twice if there were not enough different words per cell (i.e. 5, see Table 4).
3) The frequent words *ja* 'yes', *dat* 'that' and *maar* 'but' were excluded.
4) No /a:/s or /ɑ/s were selected which were preceded or followed by a vowel (e.g. *indiaan* 'indian', *chaos* 'chaos') or when the previous or following sound was unclear.
5) Foreign words and names like *hardcore* and *Jari Litmanen* (a soccer player) were excluded.
6) Words in which the vowel is followed by an ambisyllabic consonant[3] were not selected: e.g. *pakken* 'to fetch', *grappig* 'funny', *strakker* 'more tight', *plaatje* 'picture', *vangen* 'to catch', *lachen* 'to laugh', *vanaf* 'from'.
7) Sounds with an orthographical <a> which according to the standard norms can be pronounced either with [a] or with [ɑ] were excluded (e.g. the first <a> in *Marokkaan* /marɔkan/ or /mɑrɔkan/ 'Moroccan').
8) Words which were read from for example (news)papers and magazines were not taken into account.
9) Words which were uttered in an (intentionally) conspicuous way were not selected either. For the most part these were imitations.

4.3.3 Variants and coding

Table 5 Coding schemes for (a) duration and (b) place of articulation

(a) Duration		(b) Place of Articulation		
0	Short	0	ə	Unspecified
1	In-between	1	æ, ɛ	[-back, -low]
2	Long	2	a	[-back, +low]
		3	ɑ	[+back, +low]
		4	ɔ	[+back, -low]

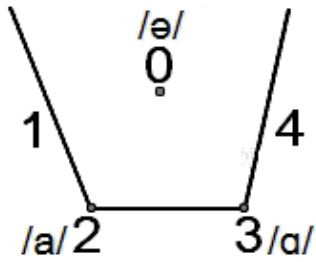


Figure 1 Coding scheme for place of articulation

Table 5 contains the variants distinguished in coding the realizations of the /a:/ and /ɑ/ variants for duration (three values or codings) and place of articulation (five values or codings). The Standard Dutch /a:/ has code 2 for duration as well as for Place of

[3] In these cases, phonetic and phonological syllable boundaries do not always coincide.

Articulation, while the Standard Dutch /ɑ/ has code 0 for duration and code 3 for PoA.

In total 3855 tokens were coded by two raters. Three-quarters of them were coded by the first author, while the fourth quarter was coded by the other rater. The agreement between the raters was evaluated by drawing a random sample of two /a:/ tokens en two /ɑ/ tokens from 20 speakers. All tokens had primary word accent and were all followed by an obstruent. All 80 tokens were scored independently by two raters. All /a:/ tokens got a score between 1 and 3 for place of articulation. The variation in scores for /ɑ/ was larger. All five variants for place of articulation turned out to occur, the most frequent one being by far the score of 3 (33 times for rater 1, 32 times for rater 2). The two raters disagreed on the assignment of the reduced variant scored by a '0' for PoA. They agreed twice, but scored each three other reduced variants. The reduced variants will not be taken into account in the analyses below, as the analysis of the type of distinction between the two vowels and their quality is not relevant in relation to the more general processes of vowel reduction in Dutch in spontaneous speech. The analyses on agreement between the two raters were done on the remaining 72 realizations. The agreement score (kappa) for place of articulation is .621, that can be qualified as acceptable. The agreement for duration was .787, indicating that the agreement was substantial.

The rater scores were evaluated by comparing them to acoustic measurements as well to investigate how valid they were. The length of the 72 vowels was measured acoustically. It turned out to be strongly correlated with the rater scores (for rater 1: .721, for rater 2: .739). We measured formant 1 (F1) and formant 2 (F2) values of the 72 vowels at 50% of the vowel length. The two formants turned out to be strong predictors of the place of articulation scores for both raters. Applying a multinomial regression analysis on rater 1 returned 81.9% correctly classified scores; the success rate for rater 2 was 84.7%.

The conclusion is that the scores of both raters are reliable and valid and can be used to estimate the properties of the variants distinguished, both which respect to duration as to place of articulation.

4.3.4 Data analyses

We first excluded all reduced (= 0) codings from the analysis. We then proceeded by analyzing the patterns of variation before an obstruent. We decided to leave out the realization before liquids because of complications in the patterns of variation in the realization of a following liquid, as many liquid variants turned out to be present, with strong differences between Amsterdam and Nijmegen. /r/ in Nijmegen is strongly reduced and uvular, whereas /r/ in Amsterdam is varying between an alveolar, uvular and bunched realization (cf. Sebregts 2015). A following /l/ turned out to be involved

in patterns of vocalization, with obvious distinctions between Amsterdam and Nijmegen.

4.4 Results

4.4.1 Overall results

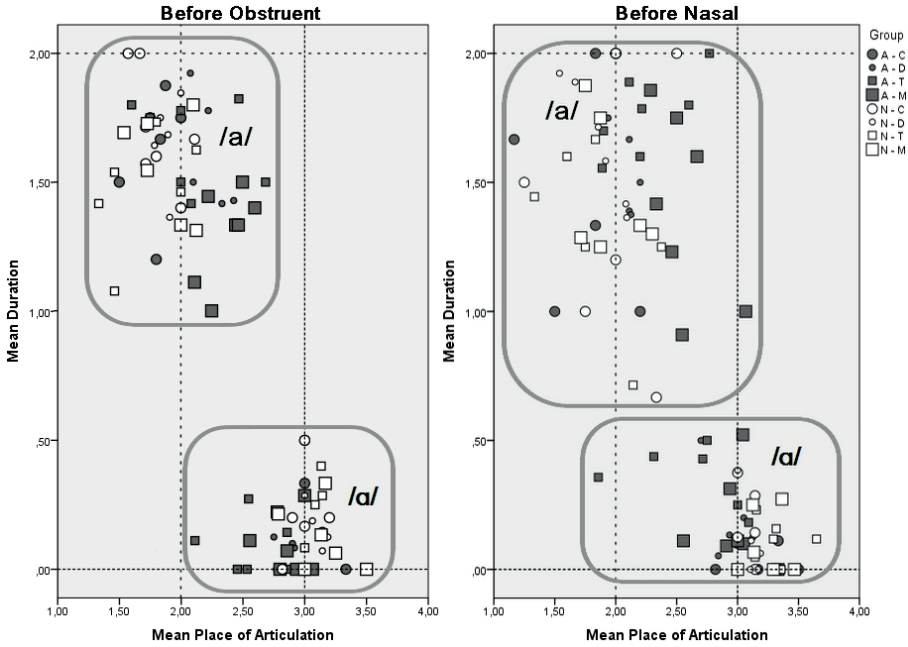


Figure 2 The realization of /a:/ and /ɑ/ in two contexts, before obstruents (left panel) and before nasals (right panel). Each symbol (dot; square) represents a speaker (some symbols overlap). A = Amsterdam, N = Nijmegen. The C and D speakers are the white Dutch speakers without inter-ethnic ties (C) and with strong inter-ethnic ties (D). The T speakers are Turkish-Dutch; the M speakers Moroccan-Dutch. The C groups only engaged in in-group conversations; the means in the other groups resemble the average of both in- and out-group conversations. The dashed lines represent a typically standard Dutch /a:/ (wide spaced dashed lines) and /ɑ/ (small spaced dashed lines) according to our coding scheme.

Figure 2 presents the mean values of /a:/ and /ɑ/ per single speaker, in two contexts. The non-liquids are split up for a following obstruent (left panel) or nasal (right panel). For each single speaker, the means were calculated over the various conversations studied (i.e. irrespective of the background of interlocutor). The place of articulation is defined by the horizontal axis, duration being defined by the vertical axis. Figure 2 shows straightforwardly that the two phonemes are marked by a robust distinction in duration in the obstruent context. There is large gap between the two phonemes (.50). Nasals give a less outspoken pattern, but there is no overlap in duration,

indicating that /a:/ is shorter on average in nasal than in the obstruent context. The gap between the two vowels is also less outspoken for place of articulation (PoA), the amount of variation and overlap again being larger in the nasal context.

We can observe for both vowels in both panels that the Amsterdam speakers (gray symbols) are situated differently on the horizontal axis. They are more to the right for /a:/ (having more back realizations than the Nijmegen speakers (white symbols)) and more to the left for /ɑ/ (having more fronted realizations than the Nijmegen speakers), resulting in an overlap in PoA. The distinction for the Nijmegen speakers is much more outspoken, with no overlapping mean scores between both phonemes.

Table 6 Significant effects for Duration and PoA in the contexts before obstruent and before nasal for the D, T, and M groups (speaker background); the PES (partial eta squared) is given for significant effects. I= interlocutor, S = speaker

	Duration				PoA			
	/ɑ/		/a:/		/ɑ/		/a:/	
	Obs	Nas	Obs	Nas	Obs	Nas	Obs	Nas
Main effects								
City		.358			.470	.455	.598	,362
Age		.251					.186	
Interaction effects								
Background S * City				.305		.245		
Background S * Background I			,274					
Background S * Background I * Age		.283				.198	.320	
Background I * City * Age				.166				

Table 7 Significant effects for Duration and PoA in the contexts before obstruent and before nasal for the C and D groups (speaker background); the PES (partial eta squared) is given for significant effects

	Duration				PoA			
	/ɑ/		/a:/		/ɑ/		/a:/	
	Obs	Nas	Obs	Nas	Obs	Nas	Obs	Nas
Main effects								
Speaker background							.388	
Interaction effects								
Speaker background * Age		.323						
Speaker background * City * Age				.214				

We tested the social contrasts in the D, T, and M groups for duration and PoA statistically by ANOVAs (GLM, Repeated Measures). In these analyses, we included the background of the interlocutor as a factor as well. Unfortunately, because of small numbers of realizations, this causes one speaker in the obstruent context to be left out in the analyses of /ɑ/, as well as one other speaker in the analyses of /a:/. In the

nasal context, this causes three speakers to be left out in the analyses of /ɑ/, as well as two other speakers in the analyses of /a:/. For these speakers too few relevant data were available for one of the three interlocutor conditions. Table 6 gives an overview of the significant effects found. We analyzed the C and D groups in in-group conversations separately from the D, T and M groups in both in- and out-group conversations. The results for the D, T, and M groups are given in Table 6, those for the C and D groups in Table 7. The Partial Eta Squared (PES)[4] is given for each significant effect. These effects will be discussed further in Sections 4.2 (duration) and 4.3 (PoA).

4.4.2 Duration

No significant extra-linguistic effect was found for *duration* of /ɑ/ in the obstruent context. The nasal context produced more results. On average, Amsterdam speakers have a higher mean duration of /ɑ/s before nasals (mean value: .27) than Nijmegen speakers (.10). This causes a significant effect of city ($F(1,22)=12.293$, $p=.002$, $PES=.358$). There is also an age effect ($F(1,22)=7.378$, $p=.013$, $PES=.251$); the 12 year olds use on average longer /ɑ/s before nasals (.24) than the 20 year olds (.13). In addition, there is a 3-way interaction effect: speaker background * interlocutor background * age ($F(4,44)=4.339$, $p=.005$, $PES=.283$). Separate analyses on the three speaker groups (D, T, M) reveals an interaction effect of age and the background of the interlocutor for the D group: $F(1.904,13.327)=4.190$, $p=.040$, $PES=.374$. The 12 year olds have longer realization in interaction with Turkish interlocutors, whereas the same effect is found for the 20 year olds but now in interaction with Moroccan interlocutors. Such an interlocutor effect was found neither for the Turkish-Dutch nor for the Moroccan-Dutch speakers.

For the duration of /a:/ before obstruents, no main effects were found. Only one interaction effect was found in the obstruent context, namely between the background of the speaker and that of the interlocutor ($F(4,48)=4.519$, $p=.004$, $PES=.274$). A first look at the patterns (nine means) of the three groups (D, T, M) in interaction with participants from these three groups (D, T, M) might suggest that the 'white' Dutch speakers with inter-ethnic ties use on average shorter /a:/'s when speaking to Turkish-Dutch speakers (1.25) than in conversation with speakers of the other two groups (D: 1.67; M: 1.79), although this difference is not significant in a post-hoc analysis. The Turkish-Dutch reply to that by using on average shorter /a:/'s when speaking to the D group (1.28) compared to speaking to the Moroccan-Dutch (1.62) and their own group (1.63), an effect that turned out to be significant

[4] A measure of effect size, comparable to the proportion of explained variance: PES is calculated after having removed the variance explained by the other variables in the analysis.

(F(2,16)=3.783, p=.045, PES=.321) in a subsequent analysis of the Turkish-Dutch group. The Moroccan-Dutch use, on the other hand, longer /a:/s when speaking to the D group (1.71) than in conversation with the Turkish-Dutch (1.37) or members of their own group (1.36), and this effect of interlocutor is significant (F(2,18)=5.4843, p=.014, PES=.379), in a subsequent analysis on the Moroccan-Dutch group. In this pattern of interactionally conditioned variation, the Moroccan-Dutch replicate the pattern of the variation in the duration of /a:/ in the speech of the members of the respective groups: M ≤ T < D. This is therefore a clear instance of a language style as audience design effect, as first discussed by Bell (1984).

Just as for /a:/ before obstruents, no main effects were found in the mean duration of /a:/ before nasals. However, two interaction effects were encountered: background of the speaker * city (F(2,23)=5.038, p=.015, PES=.305), and background of the interlocutor * city * age (F(2,46)=4.565, p=.016, PES=.166). The two-way interaction can be explained by separate analyses of Amsterdam and Nijmegen which showed no effect of background of the speaker for Nijmegen, but a significant effect for Amsterdam. The Turkish-Dutch speakers from Amsterdam (mean: 1.83) differ significantly from the Moroccan-Dutch speakers from the same city (1.34), F(2,13)=5.769, p=.016, PES=.470. The 'white' Dutch speakers with inter-ethnic ties (1.51) pattern with both groups. Another way to explain the two-way interaction (background of the speaker * city) is by analyzing the three background groups separately. The Turkish speakers from Amsterdam use significantly longer /a:/s than the ones from Nijmegen. Given the patterns for the M and D groups, the Turkish speakers from Amsterdam seem to take a distinct position with respect to the duration of /a:/ before nasals compared to the other speakers.

The three-way interaction, background of the interlocutor * city * age, is brought about by the fact that an interaction effect was encountered for Nijmegen: interlocutor background * age, F(1.820,18.198) =4.356, p=.031, PES=.303. The 12 year olds from Nijmegen use longer mean /a:/s (1.78) than the 20 year olds (1.19) when talking with 'white' Dutch interlocutors, while they use shorter mean /a:/s when speaking with Moroccan-Dutch interlocutors compared to the 20 year olds (respectively 1.16 and 1.43). There is no difference between the two age groups when talking to Turkish-Dutch interlocutors (respectively 1.63 and 1.64).

How different are the two 'white' Dutch groups (see Table 7)? When comparing /ɑ/ before nasals of the 'white' Dutch with interethnic ties (the D group) with that of the 'white' Dutch without inter-ethnic ties (the C group), no significant main effect was found between the two groups (i.e. background of the speaker). However, for duration a significant interaction effect was found between background of the speaker and age (F(1,17)=8.123, p=.011, PES=.323). The 20 year old 'white' Dutch

without inter-ethnic ties (the C group) use slightly longer /ɑ/s before nasals (.13) than the 12 year olds (.03), whereas the pattern is opposite for the 'white' Dutch with inter-ethnic ties (i.e. D group): 12 year olds .25 and 20 year olds .00.

When comparing the two 'white' Dutch groups on the duration of /a:/, no significant main effect was found between the two groups (i.e. background of the speaker). However, for /a:/ before nasals, one significant interaction effect was found in in-group situations for *duration* of /a:/ before nasals: background of the speaker * city * age, F(1,17)=4.630, p=.046. Among the Amsterdam 12 year olds, the C group members' /a:/'s are longer than the D group members' /a:/'s , while for the 20 year olds the pattern is reversed. For the Nijmegen speakers, the reverse picture obtains.

4.4.3 Place of Articulation (PoA)

We will first consider the distribution over the M, T and D groups. Examining the average PoA of /ɑ/ in the obstruent context, a clear main effect was found for city (F(1,24)=21.311, p=.000, PES=.470). The speakers from Nijmegen have a mean of 3.13, while the ones from Amsterdam have a mean of 2.71. This implies that /ɑ/ before obstruents from the Nijmegen speakers goes in the direction of [+back, -low], while that of the Amsterdam speakers goes in the direction of [-back, +low]. This is visible in Figure 2: most Nijmegen speakers are located on the right side of the vertical dashed line, while the Amsterdam speakers are mostly situated at the left.

The ANOVA of /ɑ/ in the nasal context showed a main effect of city too (F(1,22)=18.371, p=.000). The pattern is the same as for the obstruents: /ɑ/s realized by speakers from Nijmegen had a higher mean (3.24) than those realized by the speakers from Amsterdam (2.80). However, the variable city interacts with the background of the speaker significantly, F(2,22)=3.579, p=.045., PES=.245. For all three backgrounds, the Nijmegen speakers have a higher mean than the Amsterdam speakers, but the difference is the largest for the Turkish-Dutch group who score the lowest in Amsterdam (2.58 vs. 2.93 for Moroccan-Dutch and 2.87 for 'white' Dutch) and the highest in Nijmegen (3.37 vs. 3.24 for Moroccan-Dutch and 3.11 for 'white' Dutch). Just as for duration of /ɑ/ before nasals the 3-way interaction speaker background * interlocutor background * age showed up for PoA of /ɑ/ (F(4,44)=2.724, p=.041, PES=.198). No main effects and no interaction effect of age and interlocutor background were found in the separate analyses of the D, T and M group. It is clear that an interlocutor effect plays a role, but it is hard to define it more precisely.

For PoA of /a:/ before obstruents, a significant main effect of City was found (F(1,24)=35.705, p=.000, PES=.598). The Amsterdam speakers had a higher value (2.24) than the Nijmegen speakers (1.73). Also an age effect was found, F(1,24)=5.470, p=.028, PES=.186. The 20 year olds had a higher value (2.10) than

then the 12 year olds (1.88). Neither the background of the speaker nor the background of the interlocutor has a significant main effect. One significant interaction effect was found: speaker background * interlocutor background * age, F(4,48)=5.645, p=.001, PES=.320. To reach a better understanding this three-way interaction, we examined the three speaker groups (D, T and M, i.e. speaker background) separately. No relevant effects were found for the D and M speaker groups. However, a significant interaction effect between interlocutor background and age was found for the Turkish-Dutch speakers, F(2,16)=9.843, p=.002, PES=.552. The 20-year-old Turkish-Dutch speakers use more backed /a:/'s when speaking to Dutch interlocutors (mean: 2.43), than when speaking to their own peers (2.07) or to Moroccan-Dutch interlocutors (1.72). On the contrary, the 12-year-old Turkish-Dutch speakers use the more fronted /a:/'s when speaking to Dutch interlocutors (mean: 1.47) compared to speaking to Moroccan-Dutch interlocutors (1.78) and to their own peers (1.86).

Just as PoA for /a:/ before obstruents, a similar significant effect for city (F(1,23)=13.044, p=.001) was found for the nasal context. The Amsterdam speakers had a higher mean value (2.30) than the Nijmegen speakers (1.90). In contrast to the context 'before obstruent', no age effect was found nor the 3-way interaction-effect.

How different are the two 'white' Dutch groups (see Table 7)? Examining the PoA of /ɑ/ in connection with the speakers' backgrounds, we did not find any differences between the 'white' Dutch without and with inter-ethnic ties (i.e. C group and D group, respectively) in in-group conversations in both contexts ('before obstruent' and 'before nasal').

When comparing the PoA of /a:/ of the 'white' Dutch with interethnic ties (the D group) with those of the 'white' Dutch without inter-ethnic ties (the C group), a significant difference was found in the obstruent context between the two groups (i.e. background of the speaker), F(1,17)= 10.758, p=.004, PES=.388. The Dutch with inter-ethnic ties (D) had a higher mean in in-group conversations (2.09) than the Dutch without inter-ethnic ties (C, 1.81).

4.4.4 Distances between /a:/ and /ɑ/

For both duration and PoA, we calculated the distances between the two vowels by using Ashman's *D* (see Labov 2014:10). For a given vowel pair, it measures the

absolute[5] mean difference ($|\mu_1 - \mu_2|$) divided by the standard deviation calculated on the basis of the variances in both vowels. We calculated Ashman's *D* scores for each speaker for duration and PoA between /a:/ and /ɑ/.

The factor background of the interlocutor was not taken into account in the analyses, as (1) there were not sufficient data for two 20 year olds in the obstruent context and for five 12-years old speakers in the nasal context data, and (2) this factor has been found to be involved merely in certain three-way effects, and does not have any autonomous influence (see previous sections). Therefore, there is no need to analyze the data for the 'white' Dutch speakers without inter-ethnic ties (C group) separately.

Table 8 Significant effects for PoA of Ashman's *D* in the contexts before obstruent and before nasal for the C, D, T, and M groups (speaker background); PES is given for significant effects

	Obstruents	Nasals
Main effects		
City	.556	.345
Age		.123
Speaker background	.533	.301
Significant interaction effects		
Speaker background * City	.523	.317
Speaker background * City * Age	.203	

No significant effects were established for duration. The statistical results for PoA are summarized in Table 8. The ANOVAs (GLM, Univariate) analyses performed on Ashman's *D* in the obstruent context showed significant main effects for city ($F(1,35)=43.875$, $p=.000$, PES=.556), and speaker background ($F(3,35)=13.302$, $p=.000$, PES=.533). The mean value of Ashman's *D* of Nijmegen speakers (2.00) is larger than that of the speakers from Amsterdam (1.57). The main effect of speaker background has to be interpreted using the interaction effect of speaker background * city ($F(3,35)=12.771$, $p=.000$, PES=.523). The mean Ashman's *D* scores split out for groups and city are given in Figure 3, with the left panel for the obstruent context. Nijmegen is distinct from Amsterdam, as no differences show up between the speaker groups. The situation is obvious in Amsterdam. The 'white' Dutch Amsterdam speakers without inter-ethnic ties (C group) differ significantly from the other three groups.

[5] Ashman's *D* uses absolute differences, which means negative values are 'lost'. With regard to Place of Articulation, a negative value implies that the mean articulation of /ɑ/ is more fronted than that of /a:/. In the obstruent context, one speaker has a negative value: -.04. The mean /ɑ/ of this 20 years old Moroccan Dutch speaker from Amsterdam is more fronted than his mean /a:/. The same is the case for two 20 years old Turkish-Dutch speakers from Amsterdam in the nasal context: -.36 and -.05.

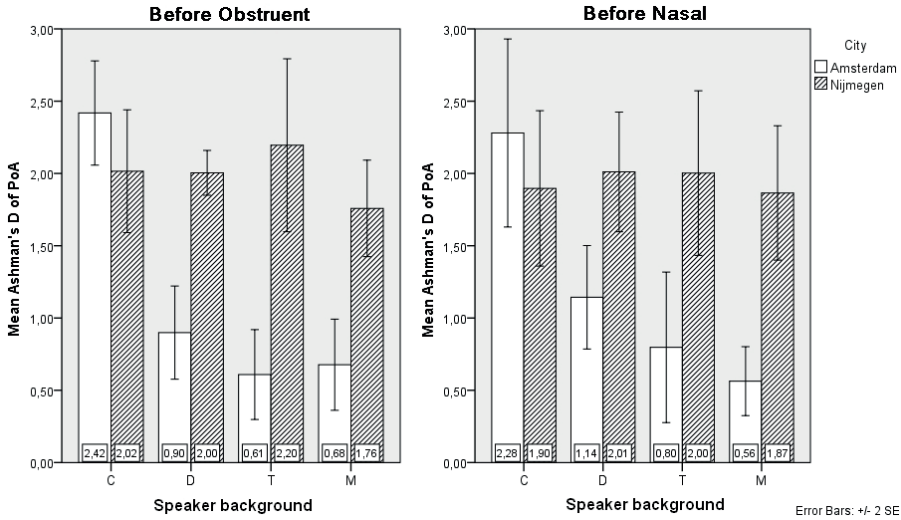


Figure 3 Mean Ashman's *D* of Place of Articulation. C = white Dutch speakers with no inter-ethnic ties, D = white Dutch speakers with strong inter-ethnic ties, T = Turkish-Dutch; M = Moroccan-Dutch

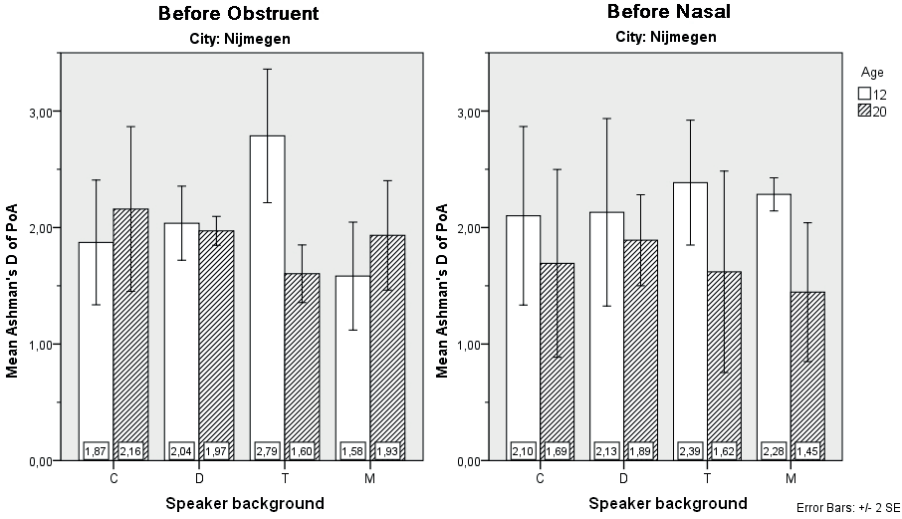


Figure 4 Mean Ashman's *D* of Place of Articulation for speakers of Nijmegen. 12 = 10-12 year olds, 20 = 18-20 year olds, C = white Dutch speakers with no inter-ethnic ties, D = white Dutch speakers with strong inter-ethnic ties, T = Turkish-Dutch; M = Moroccan-Dutch

A three-way interaction was found between the background of the speaker, city and age ($F(3,35)=2.964$, $p=.045$, PES=.203). This interaction effect can be ascribed to the fact that an interaction effect between the background of the speaker and age was found for Nijmegen. This can be seen in the left panel of Figure 4. The 12 year

old Turkish speakers in Nijmegen have on average a larger distance (2.79) between /a:/ and /ɑ/ than the 20 year old Turkish speakers (1.60) or any of the other Nijmegen speakers (group means between 1.58 and 2.16).

In the nasal context, we found comparable results as for the obstruent context, i.e. a main effect for city with a larger mean value for Nijmegen than for Amsterdam ($F(1,35)=18.419$, $p=.000$, PES=.345) and a main effect of the background of the speaker ($F(1,35)=4.928$, $p=.033$, PES=.123). The interaction effect of speaker background * city is significant as well ($F(3,35)=5.415$, $p=.004$, PES.=.317). Figure 3 makes clear that the patterns for the nasal context are similar as in the obstruent context. There are no differences between the groups in Nijmegen, but in Amsterdam the C groups stands out. There is also a main effect for age ($F(1,35)=4.928$, $p=.033$, PES=.123). The distances are smaller for the 20 years old, as can be seen for the Nijmegen speakers in Figure 4.

4.5 Conclusion, discussion and issues for future research

As far as the variation in the realization of each of the two segments is concerned, the main findings can be summarized as follows. No main effects of the background of the speaker or the interlocutor were found for the two phonemes, neither in the two linguistic conditions nor for the two dimensions of variation. Interestingly, however, the backgrounds of the speaker and of the interlocutor were involved in several interaction effects. The clearest one is a language style as audience design effect in the speech of the Moroccan Dutch speakers in the context before obstruents. Turkish Dutch speakers produce the shortest /a:/'s when speaking to the endogenous 'white' Dutch, whereas Moroccan Dutch speakers produce the longest /a:/'s when speaking to the endogenous 'white' Dutch. For both groups, whilst speaking with members of the other ethnic minority group, there is hardly any difference in /a:/ length.

For the Place of Articulation of /ɑ/, Amsterdam speakers had a significantly lower mean value (i.e. more front realizations) than the speakers from Nijmegen, while for /a:/ the opposite was established. Amsterdam speakers also produced significantly longer /ɑ/'s before nasals. The 12 year olds use on average longer /ɑ/'s before nasals than the 20 year olds; for the place of articulation of /a:/ before obstruents, the 12 year olds have lower values (i.e. more front realizations) than the 20 year olds.

Overall, more variation occurred in place of articulation than in duration. The /a:/-/ɑ/ contrast is phonetically realized by all groups, though with variation across groups in place of articulation.

As far as the variation in the differentiation between two segments is concerned, the main findings are: the young men with Moroccan and Turkish backgrounds make the same length difference as their 'white' Amsterdam and Nijmegen peers, and in this respect the Nijmegen and Amsterdam speakers at large as well as the 12 year olds and the 20 year olds are not significantly different.

The analyses of the differentiation between place of articulation of /ɑ/ and /a:/ render an essentially different picture. The Amsterdam speakers as a group make a smaller difference between the vowels than the Nijmegen speakers. In Amsterdam, the Moroccan Dutch, the Turkish Dutch and the 'white' Dutch young men with strong interethnic ties realize smaller differences between /ɑ/ and /a:/ than the 'white' young men with weak or no interethnic ties – but in Nijmegen there is no difference between the four groups of speakers. In other words, the C group in Amsterdam shows distinctive behavior which perhaps indicates that they are converging to the national standard while the other groups are maintaining the local feature. Conversely, before a nasal the 12 year old Nijmegen speakers make more pronounced differences between the two vowels than their 20 year old fellow Nijmegen citizens, whereas in Amsterdam both age groups do not differ in this respect.

In short, in both cities the Moroccan Dutch, the Turkish Dutch and the 'white' Dutch groups behave alike. There is one exception: the C group from Amsterdam deviates with respect to place of articulation. We can thus establish that the expected L2 effect (induced by the typological distinction between syllable- versus stress-timed systems between Dutch on the one hand and the heritage languages on the other) does not manifest itself.

With respect to the position of the variation in the realization of /ɑ/ and /a:/ in the speakers' verbal repertoires, we can establish that there are several instances of situationally conditioned variation. One of these is a clear style as audience design effect and it concerns the variation in the duration of /a:/ before obstruents in the speech of the Moroccan Dutch.

Finally, we want to put forward a few observations with respect to the social spread of the patterns of variation. Despite the style as audience design effect on the length of /a:/ before obstruents in the speech of the Moroccan Dutch, and despite the fact that before nasals the Moroccan Dutch produce shorter /a:/s than the Turkish Dutch, both groups do not differ in the way they differentiate between /ɑ/ and /a:/. This does not always hold for the 'white' Dutch speakers with and those without Moroccan and Turkish friends; especially before obstruents the 'white' Dutch speakers with inter-ethnic ties pattern with their Moroccan and Turkish peers.

Follow-up research of the 'Roots of Ethnolects' data could focus on other tense/lax vowel contrasts in Dutch, such as /ɔ/ versus /o:/ and /ɪ/ versus /e:/. Future work should also target other potentially variable features in the complex defined by the typological distinction between stress-timed and syllable-timed systems. Our data contain e.g. instances of vowel apocope and of the use of full forms of small and prosodically recessive but grammatically weighty function words (such as *er* and *het*, lit. 'there' and 'it', as [ɛr] and [hɛt]) in contexts where 'white' native speakers typically use reduced or cliticized forms, such as [dər] or [ər] and [ət] or [t].

5 Co-variation and varieties in modern Dutch ethnolects

Edited from: van Meel, Linda, Frans Hinskens & Roeland van Hout. 2015. Co-variation and varieties in modern Dutch ethnolects. *Lingua* [Special Issue: Frans Hinskens & Gregory R. Guy (eds.), *Coherence, covariation and bricolage. Various approaches to the systematicity of language variation]*. doi:10.1016/j.lingua.2015.10.013.

Abstract

Ethnolect features typically have different origins. In emerging ethnolects, features are moreover in flux and structural relations between variable phenomena have not yet fully crystallized, so that the strict co-occurrence, conjunction or disjunction between variants is probably rare. In this contribution we focus on the co-variation of a range of linguistic variables in emerging Moroccan and Turkish varieties of Dutch spoken in the Netherlands. We address the question whether features with different origins can be freely and randomly mixed. Is the variation entirely free and consequently co-variation as well, or are there co-occurrence restrictions on their use? When correlated usages are encountered, are they better understood as consequences of internal factors, or as indicators of social (specifically ethnic) coherence? In our data for young Moroccan and Turkish varieties of Dutch, both linguistic and social or ethnographic factors make the linguistic variables cohere, although the linguistic rhyme and reason is the first one to catch the eye. On a more refined level of analysis, one cluster of features shows no social differentiation whatsoever, while one cluster of features appears to be areally defined and two others by the speakers' ethnic background in interaction with both their age and areal location.

5.1 Co-variation: logical, statistical and structural aspects

In most socio-dialectological studies, the emphasis lies on separate, individual linguistic variables and their variants, sometimes including intermediate variants and hyperdialectisms. Much less attention has been paid to language varieties as a whole, i.e. at the level of more or less coherent language systems.

With respect to sets of variable phenomena ('linguistic variables') in specific linguistic systems some of the main questions are: (a) in which ways can linguistic variables cohere? (b) what does that mean in terms of the organization of linguistic variation and, more generally, of linguistic competence? Are the varieties of language that are commonly referred to as standard languages, vernaculars, speech styles, dialects, ethnolects, etc. coherent objects or diffuse abstractions? They are typically characterized in terms of clusters of linguistic elements: entire grammars and lexicons in the case of languages and dialects, or sets of linguistic variables in the case of sociolects, ethnolects and speech styles.

If speakers are using the available linguistic resources randomly or if they are doing relatively unconstrained 'bricolage', i.e. if they actively and idiosyncratically select from a palette of variants available in their communities of practice to construct identities, stances, and styles (Eckert 2008a), varieties are fluid. In that case the

separate variables, which may have subtly distinctive social meanings ('indexicalities'), will not co-vary and show zero to low patterns of correlation; cf. Guy (2013). If varieties are coherent, the variables associated with them should co-vary in the usage of individuals of groups of speakers.

Co-variation is a non-accidental relation between two or more variable phenomena in the language use of a speaker or members of a specific (geographical, social or cultural) group. Statistically this relation manifests itself as a correlation, i.e. the situation where the occurrence of a phenomenon *x* systematically increases the probability of occurrence of a phenomenon *y* – or, contrarily, where the occurrence of a phenomenon *x* systematically lowers the chances of occurrence of a phenomenon *y* in an utterance, in a conversational turn, in a narrative, in a community grammar etc. The latter is generally the case for two (or more) phenomena which belong to extremely different style levels, as in e.g. the case of the use of the coronal variant [ɪn] of the (ing) variable in many varieties of English on the one hand and the heavy use of impersonal constructions (including passives) common in academic prose. In such cases of systematic positive or negative relations, co-variation is statistical in the sense of: not categorical. In general, stronger correlations should indicate greater levels of lectal coherence between phenomena.

The nature of this type of relation can vary across different sectors of a speech community. The relation can be motivated internally by structural relations. Two or more morpho-syntactic phenomena can e.g. be brought about by the same parametric change; several processes of vowel change can be part of the same chain shift. A correlation between several variable phenomena can also be extra-linguistically (and sometimes *only* extra-linguistically – Becker 2015) grounded, e.g. in stylistic and/or social connections. The latter can in turn concern geographical (traditional dialects), social (social class, network, community of practice and the like) or cultural dimensions (e.g. ethnic or religious background).

In a divergent dialect, i.e. a dialect that is structurally relatively far removed from e.g. the standard variety, typically almost every single word or phrase is simultaneously marked by several dialect features. Sometimes these features vary independently from each other. An example concerns the various dialect variants of the standard Dutch past participle *gewerkt*, 'worked', in Ripuarian dialects of Dutch (spoken in the far southeast of the Dutch language area):

(1a)	ɣˈəwɪrᵊkt	'worked' past part.
(b)	ɣˈəwɪrᵊk	
(c)	jəwɪrᵊkt	
(d)	jəwɪrᵊk	

(1a-d) are all wellformed in these dialects; (1b) has undergone word-final [t] deletion (WFtD), which is a very frequently yet variably occurring feature of these dialects, (1c) shows the effect of the weakening of the voiced palato-velar fricative, [ɣ˕] → [j], which is a productive and equally variable process in these dialects; (1d) has undergone both WFtD and [ɣ˕]-weakening. These and similar cases involve several dialect features which can meet (as it were) in the realization of a given word, although in principle they vary independently of each other; yet it is often the case that they co-vary in the sense that their use is correlated, positively or negatively. This is co-variation in the sense in which the notion is commonly used, e.g. in the Anglo-American sociolinguistic literature, including Horvath & Sankoff (1987) on Sydney English, but also in studies such as Brouwer & Van Hout (1984), Van Hout (1989:247ff) for features of the Amsterdam and Nijmegen urban dialect varieties, respectively.

Categorical rather than probabilistic relations between two or more linguistic phenomena also occur; in a way, they constitute the outer limiting cases of a probabilistic relation. In one extreme case, phenomenon *x* always occurs when phenomenon *y* occurs - in a case of strict conjunction, such as implication (e.g. feeding or counter-bleeding order; Koutsoudas, Sanders & Noll 1974). For example, in Ripuarian dialects of Dutch, variable [ɣ˕]-weakening can be fed by the dialect variant /lɪɣ˕/ of the derivational suffix, the more common variant of which is /lɪk/; hence

(2) iːərlɪjə < iːəʁlɪɣ˕ə 'honest $_{INFL}$'

The form [iːərlɪkə] would bleed the weakening process; here the fricative variant of the derivation suffix feeds the weakening process (although weakening need not apply). One phenomena supports the other; applying one feature, one creates the context for application of the other feature. Therefore, conversely, in words of this formal type, [ɣ˕]-weakening implies the use of the fricative variant of the suffix – in such cases the relation between the two linguistic variables is one of logical implication. Work by Auer (1997) argues that in a similar Old World traditional dialect setting, implicational relationships of strict co-occurrence can occur between certain types of variable phenomena, motivated by structural relations among the variants.

In the other extreme scenario, phenomenon *x* never occurs when phenomenon *y* occurs, i.e. in cases of disjunction (e.g. bleeding or counter-feeding rule ordering). An example from the Ripuarian dialects of Dutch: one of the features which sets these dialects apart from most other varieties of Dutch is dorsal fricative deletion (DFD). In lexical morphemes with a rhyme consisting of a short vowel followed by a dorsal

fricative and /t/, the fricative can be deleted. As a result of compensatory vowel lengthening, non-low vowels develop a schwa offglide. In Ripuarian dialects, non-derived words such as *gedacht* 'thought' past participle, *nacht*, 'night' and *licht* 'light', constitute in principle input for both DFD and WFt-D. However, the phenomena are obviously disjunctive: they cannot simultaneously apply on the same word. In words with this structure DFD and WFtD bleed each other, i.e. one destroys the input of the other. The dialect variant of *nacht* is thus either

(3a) nɑx

or

(3b) na:t

Traditionally in Ripuarian dialects, words of this type systematically show DFD; both processes apparently apply in accordance with the 'Elsewhere condition' (Kiparsky 1973). This condition says that, whenever a given form obeys the structural description of two different rules, the more specific rule applies; in that case the more general rule is blocked – but it does apply elsewhere. Indeed, the structural description of DFD, viz. /Vçt/, forms a proper subset of that of WFt-D, viz. [[- son](t/d)]$_{phW}$. In Ripuarian dialects the latter process applies exceptionlessly elsewhere. Moreover, as a lexical rule, DFD will apply before post-lexical WFtD.[1]

In this type of cases, phenomena exclude each other and this is another type of structural dependence between linguistic phenomena. As is the case in the latter examples, this type of structural relationship (the correlation being r = –1.00) can also hold between variable phenomena.

As linguistically complex and statistically extreme types of co-variation, marked by structural dependence between elements, structures or processes, the conjunction and disjunction of otherwise variable phenomena have not been widely studied. In order to do so at all, data from relatively stable old dialects seem to be required, i.e. traditional dialects which have undergone little or no mixing with elements, structures or processes from more or less related dialects (which typically results of speaker mobility or drastic demographic shifts in the speech community). Such dialects tend to be characterized by abundant unpredictable morpho-phonological, morphological and morpho-syntactic variation (which tends to be levelled out in situations of protracted, intensive dialect contact), on top of the transparent and (often) entirely productive phonological and phonetic variation that is also typical for younger dialects. In his aforementioned paper from 1997, Auer analyses complex data from

[1] The phenomena illustrated in (1), (2) and (3) have been described and analyzed in Hinskens (1992, 1995, 1996).

traditional Lucanian dialects of Italian and from Bavarian and Alemannic dialects of German.

Coherence, as manifested in patterns of co-variation, can be less strong when no structural necessity determines the linguistic variables. Van Hout, Kruijsen & Gerritsen (2014) investigated four linguistic variables in Dutch dialects spoken along the Romance-Germanic language border in Belgium on the basis of data from 181 speakers from 13 localities. The correlations between the linguistic variables were high enough to draw the conclusion of a convincing pattern of co-variation, meaning that e.g. the degree of denasalization of French vowels in French borrowings goes along with changing the nominal gender of borrowed nouns according to the Germanic system.

Some developments lie at the crossroads between statistical co-variation and the extreme outer limiting cases of statistical relation which have grown into structural dependencies between variable phenomena. This clearly holds for the emergence of more or less separate intermediate phases in *dialect/standard continua*, such as koinés or regional varieties of the standard variety. They result from the transition from a diglossic situation (in which the national standard language, serving as the H code, and traditional dialects, serving as L codes, were kept neatly apart) into a situation with a more fluid repertoire, labelled ‘diaglossia’ by Bellmann (1998). On the resulting continua, shifting occurs rather than switching, as abrupt transitions between the constituent systems no longer exist.

The concept of intermediate varieties can be elucidated by means of the Dutch sentence "Hij heeft het aan Jan gegeven" ('He has given it to John'), which according to Hagen (1982) in eastern Brabant can be realised in at least four slightly different ways. They are, from deep dialect to "dialect free standard language" (in Hagen's transliteration):

(4a) Hij hĕggut Janne gegèève
(b) Hij hĕ ut oan Jan gegèève
(c) Hij hĕĕft ut āon Jan gegĕĕve
(d) Hij héééft ut aan Jan gegééve

In the deepest dialect realisation, (4a), the finite verb ends in the fricative /ɣ/; the indirect object is marked only through a schwa-suffix ('Janne'). Both phenomena have disappeared in the realisation in (b). Compared to (b), the realisation in (c), which represents "standard language with dialectal (accent) colouring" (thus Hagen, our translation), displays a difference with respect to the form of the finite verb, as well as in the quality of the vowel of the preposition 'aan' and of the stem part in the

past participle 'gegeven'. In connection with the two intermediate realisations (b, c) the question now is whether, for instance, the variant 'he' of the finite verb occurs more frequently with 'oan Jan' than with 'Janne', 'aon Jan', or 'aan Jan', whether it occurs more frequently with 'gegèève' than with 'gegĕĕve' or 'gegééve', and so forth. These questions concern the nature and the patterns of the co-variation between the respective variants within a specific local or regional verbal repertoire. Can local, regional, supra-regional and standard features be freely and randomly mixed in intermediate varieties? Is the variation entirely free and consequently co-variation as well, or are there co-occurrence restrictions on the use of dialect features? Are there any implicational rigid, categorical or (statistically) implicational relations? Do the phenomena co-vary (in the sense that their variation is free, although there are certain *tendential* co-occurrences or mutual blockings which may eventually change into structural relationships)? Once the correlation has reached the value |1|, there is no longer a choice. What is the nature of the relationship underlying the co-variation?

The question of how stable diaglossic intermediate varieties, notably koinés and regional standard varieties, can become, has hardly been addressed, just as little as the question how coherent their constituent dialect features can become. The few exceptions include Lameli (2004:182–203), Lenz (2003:186–192) and Kehrein (2012:227ff., 247ff., 266ff., 294ff., 309ff., 333–339), all concerning variation in specific dialect/standard continua of modern German spoken in Germany; some of the findings point at implicational relations between variable phenomena.

5.2 The research questions

In young dialects, where most features are still in flux and structural relations between variable phenomena have not yet fully crystallized, as well as in strongly levelled dialects, the strict co-occurrence, conjunction or disjunction between variants is probably rare. In young ethnolectal varieties, linguistically the situation is probably comparable to that of young dialects, which are still in the course of developing and which are thus very much in motion. For that reason and also because it will probably reveal some of the social profiling which is part of the essence of ethnolectal speech, we will study the relation between several phenomena which occur variably in two young ethnolectal varieties of modern Dutch.

The question of the coherence of the features that constitute an ethnolect is the more urgent since, unlike in home-grown endogenous varieties, the features tend to have different origins. Whereas some of the variable phenomena may result from the adoption of local or regional dialect features, others may originate in substrate effects (Bills 1977; Carlock & Wölck 1981) and yet others are rooted in L2 acquisition. The latter are typically ‘multi-ethnolectal’ features (Clyne 2000; Quist 2000).

More generally, when correlated usages are encountered, are they better understood as consequences of internal factors, or as indicators of the social (and specifically ethnic) coherence of a variety? Put differently, is coherence primarily a structural or rather a social phenomenon? Does it pertain to the linguistic system or rather to specific sectors in the speech community? Or both? If so, then how do the two interlock?

5.2.1 Co-variation and coherence in data from the Roots of Ethnolects project

In this contribution we will address the question of co-variation of linguistic variables in emerging ethnolectal varieties of Dutch spoken in the Netherlands. These new varieties have hardly been studied systematically yet. One of the exceptions is the 'Roots of Ethnolects' project.[2] For this project, speech data were collected in the Dutch cities of Amsterdam, in the north-western part of the country, and Nijmegen, situated in the southeast. The two cities are situated in different dialect areas. Audio recordings were made of 10 to 12 and 18 to 20 year old male youngsters. The sample of speakers has been balanced for ethnic background (Moroccan-Dutch, Turkish-Dutch, two of the largest ethnic minorities in the Netherlands today, as well as 'white' Dutch). All speakers were born and raised in the Netherlands and Dutch is (one of) their mother tongue(s).

For the present study we investigated a broad range of linguistic variables from the sound components (n=13) and the domain of morpho-syntax (n=4). The set of linguistic variables is rather large in order (a) to study the emergence of patterns of co-variation, leading to recognizable and distinguishable varieties such as ethnolects, (b) to be able to distinguish between processes of divergence and convergence shaping the linguistic repertoires in urbanized contexts, and (c) to balance endogenous dialect features with variable phenomena which may be more typical of ethnic groups or of stages in the linguistic socialization. The data for the variable phenomena studied (briefly sketched in Section 5.3.3 below) are quantitative in nature and they were submitted to a range of statistical techniques.

We will consider the question whether and, if so, which linguistic divisions emerge between the ethnic groups, the age groups and the cities involved. In so far as ethnicity plays a role, do the Moroccan-Dutch and the Turkish-Dutch speakers share a general 'non-native' identity, separating them from their white, endogenous Dutch peers? Which role do the endogenous dialect features play in these developments? Which roles do contact-induced and acquisition-related phenomena

[2] This study has benefited greatly from the contributions of current and previous project members. The project is financed by the Netherlands Organisation for Scientific Research (NWO), Meertens Instituut (KNAW) and Radboud University.

play? Is the emergence of ethnolects as separate varieties enhanced primarily by external (social) or rather by internal (linguistic) factors?

The dividing lines between the ethnic groups (if any) may run along fairly arbitrary subsets of linguistic variables; such processes would be instantiations of bricolage. The dividing lines may rest on weak co-variation between the linguistic features or variables investigated, i.e. zero to low correlations between the linguistic variables involved, giving room to incidental choices of linguistic variants. This could constitute the starting level of ethnolect formation, marked primarily by personal bricolage resulting in diffuse but recognizable patterns.

5.3 Methods

5.3.1 Speakers

Data were collected from 51 youngsters from three different groups: Moroccan-Dutch, Turkish-Dutch and white Dutch. The groups were controlled for age (10-12 versus 18-20 years old), residence (Amsterdam vs. Nijmegen) and, in the case of the white Dutch speakers, the presence or absence of regular contacts with Turkish-Dutch and Moroccan-Dutch friends. Table 1 summarizes the speaker design.

Table 1 Overview of the research design and the number of speakers

Background →	Moroccan-Dutch (M)		Turkish-Dutch (T)		'white' Dutch (D)		(C)	
Inter-ethnic ties? →	Yes		Yes		Yes		No	
Years of age	10-12	18-20	10-12	18-20	10-12	18-20	10-12	18-20
Amsterdam	3	4	3	4	2	3	3	3
Nijmegen	3	3	3	3	3	3	4	4

Like the white Dutch speakers, the speakers with a Moroccan or Turkish background were born in the Netherlands and have at least one parent who immigrated from Morocco or Turkey respectively. All speakers grew up in their place of residence (either Amsterdam or Nijmegen). The white Dutch speakers were split up into two groups: those with inter-ethnic ties, that is having friends with a Moroccan or Turkish background (labelled D), and those without inter-ethnic ties (labelled C).

Due to unanticipated complications during the fieldwork sessions, only scant information became available about the language skills and linguistic profiles of the speakers in our sample. The information available is detailed in Van Meel et al. (2013). To summarize, all Turkish-Dutch speakers reported knowledge of Turkish and all Moroccan-Dutch speakers reported to know either one or more of the Moroccan languages (several dialects of Moroccan Arabic and Berber). No reliable information regarding the speakers' relative proficiencies in the languages mentioned

or their varieties is available.

5.3.2 Speech data

The speakers were recorded in free conversations of about one hour. Each conversation involved two peers of the same age group (either two 10-12 year olds or two 18-20 year olds). The speakers with a white Dutch background having inter-ethnic ties (D-group) as well as those with a Moroccan-Dutch or Turkish-Dutch background were recorded in at least one in-group conversation and two different out-group conversations. In an in-group situation the conversation partner was a peer from their own ethnic group, while in the out-group situations they talked with a peer from one of the other two ethnic groups. The recordings lasted about one hour for each pairing, so for each of these speakers approximately three hours of conversational speech were recorded.

The white Dutch speakers with no inter-ethnic ties (C-group) were only recorded in in-group conversations with fellow Dutch that equally have no inter-ethnic ties. They serve as the control group for the white Dutch speakers with strong inter-ethnic ties (D-group).

The majority of the speakers attended the same school as their conversation partners, and many of them were classmates (especially so in the case of the 10-12 year olds). To ensure a more natural conversation the interviewer was only present at the beginning and the end of the conversation. However, especially so in the case of the 10-12 year olds, additional guidance was needed to keep the speakers talking for one hour (i.e. suggesting topics to talk about, introducing card games and the like).

5.3.3 Linguistic variables

In this study we investigated variation in the realisation of five different Dutch phonemes and two morpho-syntactic phenomena: the voiced coronal fricative /z/ at the beginning of prosodic words, the front unrounded diphthong /ɛi/, the lax low vowel /ɑ/, the tense low vowel /a/, the lax mid front unrounded vowel /ɛ/, and the morpho-syntactic phenomena of neuter nominal gender marking and *hun* ‘them’ as 3 pl. subject. The vocalic variables, involving /ɑ/ and /a/, /ɛi/ and /ɛ/, are among the features distinguishing the Amsterdam and Nijmegen urban dialects.

We defined one or more linguistic variables for each phenomenon, depending on the contexts of occurrence or processes involved, leading to a total of 17 linguistic variables as shown in Table 2. The scale values can be binomial (with the values 0 and 1) or they have a range of more values/integers. Each variable will be discussed

below. All variants mentioned in the table did occur in the data.

Table 2 Scales of the 17 linguistic variables

		Scale (Values + variants)	Value Traditional Standard Dutch
/z/ (at the beginning of prosodic words)	/z/ voicing after obstruent	0 = [z] 1 = [s]	1
	/z/ voicing after vowel	0 = [s] 1 = [z]	1
	/z/ voicing after sonorant	0 = [s] 1 = [z]	1
	/z/ dentalization	0 = dentalized 1 = not dentalized	1
/ɛi/	/ɛi/ monophthongization	1 = fully monoph-thongized 2 = slightly monoph-thongized 3 = clear diphthong	3
	/ɛi/ height first element	1 = [a] 2 = [æ] or [ɑ] 3 = [ɛ] 4 = [e] or [ɪ]	3
/ɑ/ and **/a/** (followed by obstruents)	/ɑ/ duration	1 = Long 2 = Medium 3 = Short	3
	/a/ duration	1 = Short 2 = Medium 3 = Long	3
	difference duration /a/ and /ɑ/	0 1 2	2
	/ɑ/ Place of Articulation (PoA)	1 = [æ] or [ɛ] 2 = [a] 3 = [ɑ] 4 = [ɔ]	3
	/a/ PoA	1 = [ɔ] 2 = [ɑ] 3 = [a] 4 = [æ] or [ɛ]	3
	difference PoA /ɑ/ and /a/	0 1 2 3	n.a.
/ɛ/ in *heb je/ik* 'have you/I'	/ɛ/ height	1 = [ɪ] 2 = [ɛ] 3 = [æ]	2
neuter gender marking	neuter gender: article	0 = *de* 1 = *het*	1
	neuter gender: demonstrative	0 = *die* 1 = *dat*	1
	neuter gender: adnominal	0 = '*een ADJ-e'* or *'onze'* 1 = '*een ADJ-ø'* or *'ons'*	1
hun 'them', subject 3 pl.	*hun* 'them', subject 3 pl.	0 = *hun* 1 = *zij* or *die* (subject 3 pl.)	1

5.3.1.1 Phonetic variables: voicing of /z/ (at the beginning of prosodic words in three contexts)

Van Meel et al. (2013) studied whether /z/ at the beginning of prosodic words is voiced or devoiced. This was done for three separate and complimentary conditions of occurrence: after obstruent, after vowel and after sonorant, giving three linguistic variables. For instances of voiced realizations it was noted if the /z/ had a (standard

Dutch) alveolar or rather dental realization.

5.3.3.2 Phonetic variables: monophthongization and height of /ɛi/

For the phoneme /ɛi/, two linguistic variables were distinguished for all occurrences preceding a plosive or a fricative: /ɛi/ monophthongization and the height of the first element of /ɛi/. The first linguistic variable addresses the degree of monophthongization and was coded on a three-point scale with 1 being a fully monophthongized /ɛi/, hence [ɛː], and 3 a clear diphthong. The second linguistic variable is the height of the first, prominent element of /ɛi/ and this was coded on a four-point scale with the value 2 representing a typical standard Dutch first element of /ɛi/. More information on both linguistic variables can be found in van Van Meel et al. (2014).

5.3.3.3 Phonetic variables: duration and Place of Articulation of /ɑ/ and /a/ (before obstruents)

In this study we incorporated six linguistic variables that deal with the phonemes /ɑ/ and /a/ when followed by obstruents.[3] For both phonemes we coded on a three-point scale the duration, one end of the scale representing short (like a standard Dutch /ɑ/) and the other one long (like a standard Dutch /a/). The place of articulation (PoA) was coded on a four-point scale with the value 3 being a standard Dutch /a/ and 3 a standard Dutch /ɑ/. Finally, differences between /ɑ/ and /a/ were calculated for both duration and PoA, giving two additional linguistic variables.

5.3.3.4 Phonetic variable: height of /ɛ/ in heb je/ik 'have you/I'

Realizations of the phoneme /ɛ/ were coded for height in the phrases *heb je/ik* 'have you/I' (in sentences like *Heb je een broertje?* 'Do you have a little brother?' and *Die heb ik nodig* 'I need that one'). A five-point scale was used: 1 = [ɪ], 2 = [ɛ], 3 = [æ], and 1.5 and 2.5 as in-between values.

5.3.3.5 Morpho-syntactic variables: marking nominal neuter gender in three word classes

Each nominal marking of neuter gender was coded for whether the speaker used a standard Dutch marking or not. This was done for three conditions: articles (e.g. standard *het woord*, non-standard *de woord* 'the word'), demonstratives (e.g.

[3] Variable /a/ (before obstruents) in this chapter differs from variable '/a:/ Obs' in Chapter 4 (Study on /a:/ and /ɑ/), in the following ways: (1) both open and closed syllables are included in variable /a/ in the co-variation study, while variable '/a:/ Obs' only includes closed syllables, (2) in the co-variation study, the scale of /a/ (cf. Table 2) is reversed with respect to the /a:/ in Chapter 4 (Cf. Table 5 of Chapter 4).

standard *dat woord*, non-standard: *die woord* 'that word'), and adnominals (adjectives, e.g. standard *een Engels woord*, non-standard *een Engelse woord* 'an English word', and possessives, 1 person, plural, e.g. standard *ons woord*, non-standard *onze woord* 'our word').

5.3.3.6 Morpho-syntactic variable: hun 'them', 3 pl. subject

For this linguistic variable (which unlike most other variables is a nominal variable) proportions were calculated per speaker by dividing the number of *hun* ('them', 3 pl. subject) realisations by the total number of *hun* + *zij* + *die* ('them' + 'they' + 'those', all 3 pl. subject) realisations.

For all these variables, indexes were calculated for each speaker. For some speakers no index could be calculated for one of more of the morpho-syntactic variables (this applies to one speaker for neuter gender: article, two for neuter gender: demonstrative, two for neuter gender: adnominal, and three for *hun* 'them', 3 pl. subject). These missing values were replaced by the mean of all other speakers of the variable in question.

5.3.4 Standardization

The indexes were recoded to a scale that runs from 0 to 100 to make sure that (1) the direction of the scales are interpretable, and (2) the scales' ranges are comparable. For 11 of the 17 linguistic variables, a score of 100 means that the (traditional) Standard Dutch variant is used in all cases. However, for the height of the first element of the diphthong /ɛi/, the height of /ɛ/, the PoA of /ɑ/ and of /a/, the Nijmegen and Amsterdam dialects deviate from the standard norm in opposite directions and in our data all urban dialect variants occur. For the 4 linguistic variables concerned, the traditional Standard Dutch variant is therefore not at the extreme end of the scale, i.e. 100; for /ɛi/ height first element, /ɑ/ PoA and /a/ PoA this variant has a score of 66.67, while for /ɛ/ height it is 50. For the linguistic variables 'Difference duration /a/ and /ɑ/' and 'Difference PoA /a/ and /ɑ/', 100 represents a maximum difference between the mean realisations of /a/ and /ɑ/, while a score of 0 implies that both phonemes were pronounced identically (based on means).

5.4 Results

5.4.1 Effects of social factors

As a first step, we investigated the influence of the social variables (background of the speaker, city and age) on the 17 linguistic variables. Table 3 gives an overview of

the significant effects found in the ANOVAs (the most important effect for each variable is indicated in **bold**).

Table 3 Significant effect sizes, PES, per linguistic variable;
B = Background of the speaker, C = City, A = Age

	B	C	A	B*A	C*A	C*B	C*B*A
/z/ voicing after obstruent	**.443**	–	–	–	.227	–	–
/z/ voicing after vowel	**.299**	–	–	.244	.214	–	–
/z/ voicing after sonorant	.328	.172	–	.365	.171	–	.201
/z/ dentalization	**.442**	–	–	–	–	–	–
/ɛi/ monophthongization	**.615**	–	.120	–	–	–	–
/ɛi/ height first element	.468	**.784**	.362	.335	–	–	.319
/ɑ/ duration	–	–	–	–	–	–	–
/a/ duration	–	–	–	–	–	–	–
difference duration /a/ and /ɑ/	–	–	–	–	–	–	–
/ɑ/ PoA	.199	**.384**	–	–	–	.336	–
/a/ PoA	.369	–	–	.360	–	–	–
difference PoA /ɑ/ and /a/	.282	**.600**	–	–	–	.448	–
/ɛ/ height	–	**.759**	.253	–	–	–	–
neuter gender: article	**.355**	.111	–	–	–	–	–
neuter gender: demonstrative	**.661**	–	.212	–	–	–	–
neuter gender: adnominal	**.570**	–	–	–	–	.258	–
hun 'them', subject 3 pl.	**.515**	–	–	–	–	–	–

The ethnic background of the speaker has an effect on 13 of the 17 linguistic variables. For 8 of these linguistic variables, it is also the most important effect, i.e. with a much higher partial eta squared (PES)[4] value than the other significant effects. These 8 linguistic variables are all four morpho-syntactic variables (3 variables concerning neuter gender marking as well as *hun* 'them', 3 pl. subject) and 4 of the phonetic variables (/z/ voicing after obstruent, /z/ voicing after vowel, /z/ dentalization, and /ɛi/ monophthongization). This does not apply to /a/ PoA, since the PES of the interaction with the speakers' age is almost as high.

City has a significant effect on 6 variables. In four of these cases (/ɛi/ height first element, /ɑ/ PoA, Difference PoA /ɑ/ and /a/, en /ɛ/ height), it is the most important factor, i.e. it has a (much) higher PES value than the other significant effects.

Age (A) exerts much less influence than speaker background (B) and city (C). It appears to be a main effect 4 times, but it is not the most important factor in any of these cases (i.e. in all 4 cases, background of the speaker and/or city plays a more important part with higher PES values). However, Age does play a part in several interaction-effects as visible in Table 3.

[4] A measure of effect size, comparable to the proportion of explained variance: PES is calculated after having removed the variance explained by the other variables in the analysis.

5.4.2 Clusters

5.4.2.1 Linguistic variable clusters

To establish which linguistic phenomena group together in their variational patterns, we performed a cluster analysis.[5] Cluster analysis does not make any assumptions about the underlying distribution of the data. Moreover, this algorithm considers how similar entities are (e.g. in terms of Euclidean distance between the entities at issue) rather than how correlated they are, which underlies factor analysis.

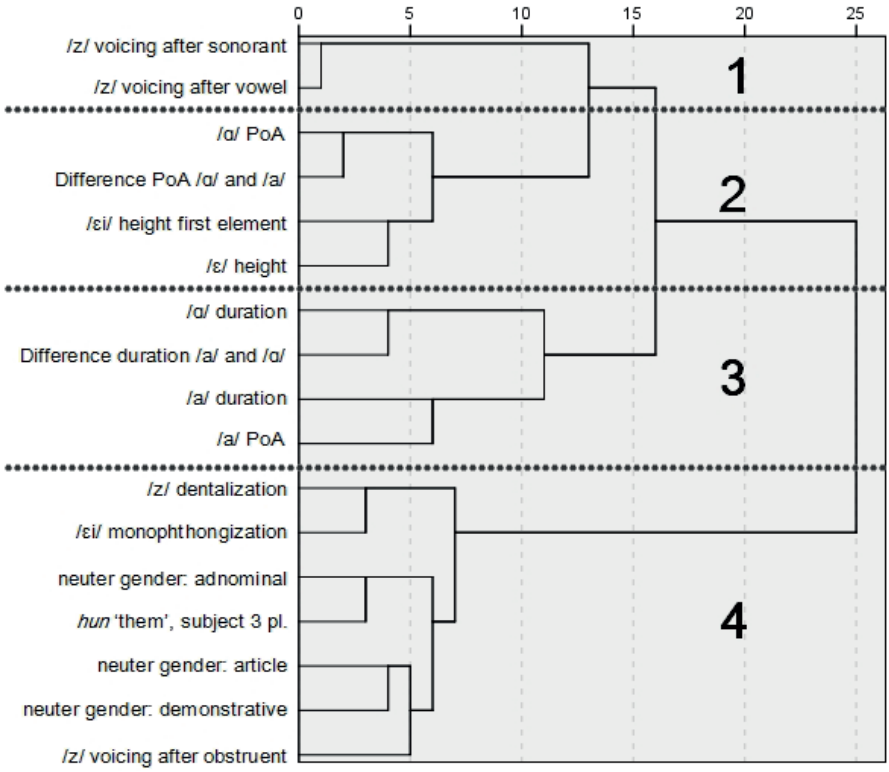


Figure 1 Dendrogram of the 17 linguistic variables, distinguishing four clusters

In order to establish which linguistic variables group together, z-scores were used to make sure that all variables are given even weight. Figure 1 shows the dendrogram with the resulting clusters. As the figure shows, four linguistic variable clusters can be distinguished.

Cluster 1 in Figure 1 contains two phonetic variables. These concern the voicing of /z/ following inherently voiced segments.

[5] We applied a hierarchical cluster analysis (SPSS), with Ward's method and Euclidean distances.

The four variables in *cluster 2* all involve phonetic variation in connection with the place of articulation (high/low, front/back) of the relevant vowels. Based on this linguistic interpretation, one would expect the PoA of vowel /a/ ('/a/ PoA') also to be included in this cluster instead of in cluster 3. However, if we look (in Table 3) at the four variables that belong to cluster 2, we notice that these are exactly the four variables for which the most important extra-linguistic factor is city. Table 4 summarizes the sizes of the significant effects (PES) of the speakers' age (A), background (B) and city (c) on each cluster of linguistic variables cluster. It confirms that city is the only single extra-linguistic factor that matters for variable cluster 2. Linguistic variable '/a/ PoA' has no significant effect of city (see Table 3).

Except for '/a/ PoA', all linguistic variables of *cluster 3* have something to do with the duration of vowels. Another uniting property is the fact that they do not show any social effect as can be seen in both Tables 3 and 4; this is an important and thought-provoking finding to which we will return below.

Table 4 Significant effects sizes (PES) per linguistic variable cluster; B = Background of the speaker, C = City, A = Age

	B	C	A	C*B	B*A	C*A	C*B*A
Variable Cluster 1 (/z/ voicing after sonorant +/z/ voicing after vowel)	.332	.124	–	–	.331	.213	–
Variable Cluster 2 (/ɑ/ PoA + diff PoA of /ɑ/ and /a/ + /ɛi/ height + /ɛ/ height)	–	**.866**	–	.445	–	–	–
Variable Cluster 3 (/ɑ/ duration + diff duration of /ɑ/ and /a/ + /a/ duration + /a/ PoA)	–	–	–	–	–	–	–
Variable Cluster 4 (/z/ dentalization + /ɛi/ monophthongization + neuter gender: adnominal + *hun* 'them', subject 3 pl.+ neuter gender: article + neuter gender: demonstrative + /z/ voicing after obstruent)	**.837**	.188	.117	–	–	–	–

Linguistic variable *cluster 4* contains all four morpho-syntactic phenomena (all in the field of inflection) as well as three phonetic phenomena which are all variables with ethnolectal variants. Both the voicing of /z/ after obstruents and the dental realisation of /z/ are exogenous, 'un-Dutch' phenomena, while monophthongization of /ɛi/ is going through a social redistribution, i.e. the monophthongal variant is being rejected by white Dutch speakers and ethnolecticized by the Turkish-Dutch and Moroccan-Dutch speakers (Van Meel et al. 2014). We saw already in Table 3 that speaker background is the most important extra-linguistic factor for these variables. Table 4

confirms that this is by far the most important factor for linguistic variable cluster 4; at the same time, speaker background does not have a stronger effect in any other variable cluster. A Tukey Post Hoc test with cluster 4 as the dependent variable makes clear that the two white Dutch groups (labeled C and D in Table 1) form one subset separate from the Turkish and Moroccan speakers (i.e. T + M) who form another subset.

5.4.2.2 Speaker clusters

We performed the same type of cluster analysis to establish which of the speakers group together.[6] On the basis of their linguistic behavior, i.e. their behavior with regard to the 17 linguistic variables, five speaker clusters can be distinguished, as can be seen in Figure 2 that contains the dendogram with the speaker clusters.

There is a clear division between the 'white' Dutch speakers (C+D) and the Turkish-Dutch and Moroccan-Dutch speakers (T+M): *speaker clusters 3, 4 and 5* contain only Turkish-Dutch and Moroccan-Dutch speakers, while *speaker cluster 1* contains only 'white' Dutch speakers. Only *speaker cluster 2* is an exception as it contains not only 'white' Dutch, but also one Turkish-Dutch and two Moroccan-Dutch speakers (indicated by the grey square in Figure 2).

A further examination of the two 'white' Dutch clusters reveals that *speaker cluster 2* only contains speakers from Nijmegen. As it belongs to the 'white' Dutch half of the speaker sample and only contains Nijmegen speakers, it is labelled 'C+D, N'. *Speaker cluster 1* is labelled 'C+D, A+N' as it contains not only all 'white' Dutch speakers from Amsterdam, but also half of the 'white' Dutch speakers from Nijmegen.

The examination of the three T+M clusters reveals that also here there is one cluster that only contains speakers from Nijmegen, i.e. *speaker cluster 5*, labelled 'T+M, N'. *Speaker cluster 4* contains mainly 12 year old speakers from Amsterdam as well as one 20 year old from Amsterdam and a 12 year old from Nijmegen. Therefore, this cluster is labelled 'T+M, A12'. *Speaker cluster 3* contains the remaining 20 year old T+M speakers from Amsterdam. It also includes two 12 year olds from Amsterdam and one 12 year old from Nijmegen. This speaker cluster is labelled 'T+M, A20'.

[6] We ran hierarchical cluster analysis (SPSS), applying Ward's method to Euclidean distances; the linguistic variables were normalized.

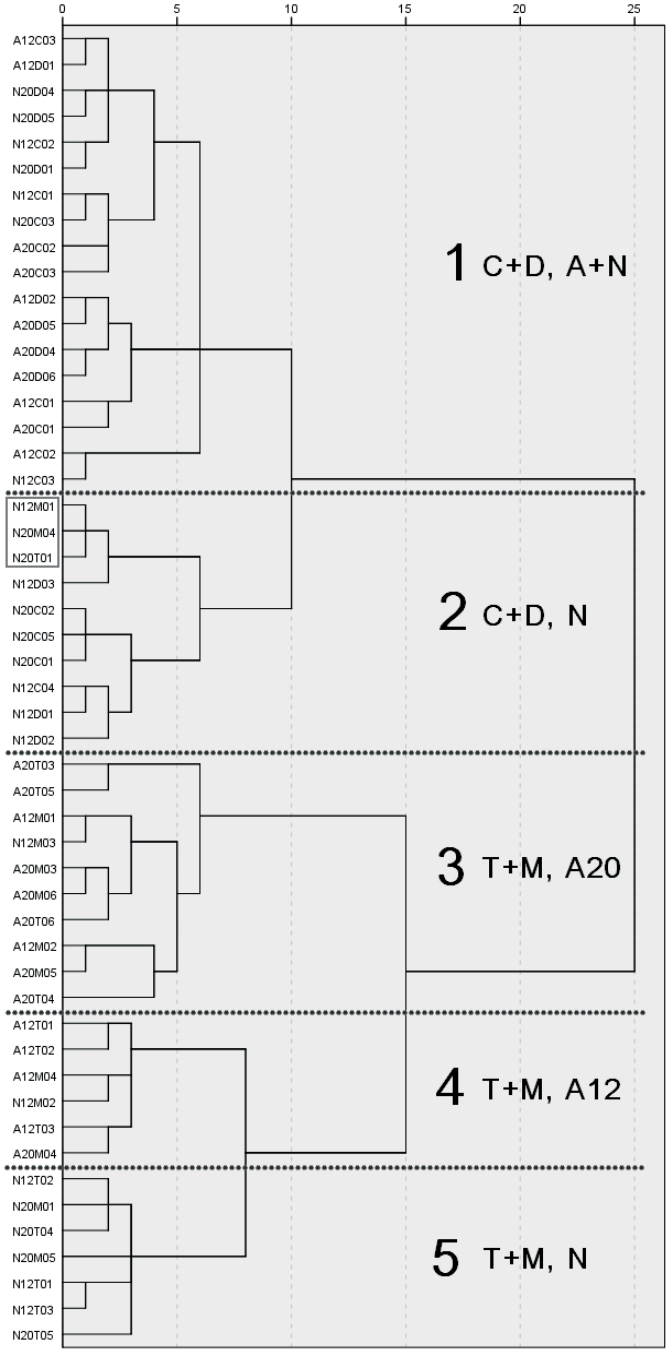


Figure 2 Dendrogram of five speaker clusters. A20 = 20 year olds from Amsterdam, A12 = 12 year olds from Amsterdam, C = white Dutch without ethnic ties, D = white Dutch with ethnic ties, M = Moroccan-Dutch, N = Nijmegen, T = Turkish-Dutch

5.4.3 Relations between speaker clusters and variable clusters

In the previous section, clusters of linguistic variables (Section 5.4.2.1) and clusters of speakers (Section 5.4.2.2) were identified. The question is how both types of clusters are related.

First of all, we needed to determine whether and, if so, how the five speaker clusters differ with respect to the use of the variable phenomena. To this end, Tukey Post Hoc analyses[7] were carried out for each linguistic variable. This analysis (a) rank orders the five speaker groups, going from the group with the lowest mean to the group with the highest mean, (b) signals which groups differ significantly from each other and (c) divide the groups in one or more subsets. The five groups did not differ significantly (i.e. form one homogenous set) in the case of two linguistic variables (duration of /a/ and of /ɑ/). These will be indicated by 0. For four linguistic variables, Tukey Post Hoc shows two clearly separated subsets of speakers (e.g. /z/ voicing after sonorants). The speaker groups in subset 1 do not differ from each other and will all be labelled with – –, while the speaker groups in subset 2 also do not differ from each other and will be labelled with ++. Five linguistic variables have a three-part marking (e.g. /ɛ/ height). The speaker groups marked with – – differ significantly from the groups marked with ++, while the speaker groups marked with +/– do not significantly differ from the – – or the ++ speaker groups. Five other linguistic variables have a four-part marking (– –, –, +, ++; e.g. /ɑ/ PoA) with ++ and – – significantly different. The speaker groups with + only differ significantly from the speaker groups with – –, while the groups with – only differ significantly from the speaker groups with ++. There is one linguistic variable with a five-part marking, viz. neuter gender: adnominal. The marking/labelling in shown in Table 5.

Table 5 Tukey Post Hoc of 'neuter gender: adnominal' with labels used in Table 6

Speaker cluster	N	Subset			Marking in Table 6
		1	2	3	
5	7	50,4286			– –
3	10	66,6000	66,6000		–
4	6		73,2381		+/–
2	10		87,7429	87,7429	+
1	18			98,1111	++
Sig.		.242	.064	.668	

For 9 of the 17 linguistic variables, the speaker cluster(s) with ++ used, relatively speaking, the most normative Standard Dutch realizations in comparison to the other

[7] In connection with ANOVA's GLM Univariate analyses. The dependent variable was formed by each of the 17 linguistic variables (one at a time); the speaker clusters were treated as a fixed factor.

speaker clusters. In other words, ++ marks the speaker cluster(s) with the closest approximation of the prescriptive standard Dutch norms for the variable at issue; e.g. for the linguistic variable /z/ dentalization, the prescriptive standard Dutch norm is: no dentalization, symbolized as ++. In our data this applies most to speakers of cluster 1, C+D A+N, i.e. the white Dutch from Amsterdam and Nijmegen. Conversely, – – marks the speakers who are most removed from the prescriptive standard Dutch norms for the variable at issue (e.g. most dentalization of /z/), in our data in the output of speaker clusters 4 and 5, the 12 year old Amsterdam as well as all Nijmegen Turkish Dutch and Moroccan Dutch speakers. In the case of the variables height of the first element of /ɛi/, PoA of /ɑ/, PoA of /a/, height of /ɛ/, and 'difference PoA /a/ and /ɑ/' it is not possible to say who is closest to or furthest away from the standard Dutch norm, as the norm is not 100, as described in Section 5.3.4.

The results of this exercise were then entered in a table in which the linguistic variable clusters and the speaker clusters are crossed; see Table 6, below. This table should help to answer the question whether and how the clusters of linguistic variables and the clusters of speakers are related. The table also reveals whether the five speaker clusters treat the linguistic variables in all four clusters consistently.

Let's first have a closer look at *linguistic variable cluster 1* that consists of two linguistic variables (phonation of /z/ after sonorant and after obstruent). A clear break is visible between speaker clusters 1, 2 and 3 on the one hand, and speaker clusters 4 and 5 on the other hand. This is confirmed by a Tukey Post Hoc analysis with the sum of the standardized scores for the linguistic variables belonging to variable cluster 1 as dependent variable and the speaker clusters as fixed factor (outcomes not discussed here). It can thus be established that the white Dutch speakers do not approximate the traditional standard norms the most (cf. Van Meel et al. 2013); in fact, the white Dutch speakers are on the far non-standard side of the scales. Table 4 (Section 5.4.2.1 above) shows that for linguistic variable cluster 1 age has a significant effect in interactions with city as well as background of the speaker. This might explain why the Amsterdam T+M speakers are separated in two different clusters with a 12-year-old cluster and a 20-year-old one.

The *second linguistic variable cluster* is largely influenced by the factor city, as we already saw in Section 5.4.2.1. This is also noticeable in the division of the speaker clusters. There is a clear difference between the Nijmegen speaker clusters 2 and 5 on the one hand and the Amsterdam speaker clusters 3 and 4 on the other hand; this patterning approaches a complementary distribution. The 'mixed' speaker cluster 1 (i.e. all Amsterdam C+D and half of the Nijmegen C+D speakers) seems to be in between those when looking at the symbols of variable cluster 2 in Table 6. A Tukey Post Hoc analysis with variable cluster 2 as dependent variable shows that

speaker clusters 2 and 5 form one subset and that speaker cluster 1 forms another subset together with the Amsterdam speaker clusters 3 and 4.

Table 6 Division of the four linguistic variable clusters by the five speaker clusters; the scores of all linguistic variables are being evaluated for the five speaker clusters; – – = speaker cluster(s) with lowest mean value(s), ++ = speaker cluster(s) with highest average value(s)

	Linguistic variable	*Speaker cluster 1* C+D, A+N	*Speaker cluster 2* C+D, N	*Speaker cluster 3* T+M, A20	*Speaker cluster 4* T+M, A12	*Speaker cluster 5* T+M, N
Variable cluster 1	/z/ voicing after sonorant	– –	– –	– –	++	++
	/z/ voicing after vowel	– –	– –	– –	++	++
Variable cluster 2	/ɑ/ PoA	+	++	– –	–	++
	difference PoA /ɑ/ and /a/	++	++	– –	– –	++
	/ɛi/ height first element	– –	+	–	– –	++
	/ɛ/ height	+/–	++	– –	– –	++
Variable cluster 3	/ɑ/ duration	0	0	0	0	0
	Difference duration /a/ and /ɑ/	+/–	++	+/–	+/–	– –
	/a/ duration	0	0	0	0	0
	/a/ PoA	+/–	– –	++	++	++
Variable cluster 4	/z/ dentalization	++	+	–	– –	– –
	/ɛi/ monophthongization	++	+	– –	–	–
	neuter gender: adnominal	++	+	–	+/–	– –
	hun 'them', 3 pl. subject	++	+	–	–	– –
	neuter gender: article	++	+/–	+/–	+/–	– –
	neuter gender: demonstrative	++	+/–	– –	– –	– –
	/z/ voicing after obstruent	++	++	++	– –	– –

In Section 5.4.2.1 it was already established that *variable cluster 3* shows no significant effects of the speakers' age, background or city. Although Table 6 shows a three-way division of speaker clusters for two of the 4 linguistic variables (i.e. 'Difference duration /a/ and /ɑ/' and /a/ PoA)[8] in variable cluster 3 (see Table 6), a Tukey Post Hoc analysis with variable cluster 3 as dependent variable confirms the

[8] In almost two thirds of the cases, the Turkish-Dutch and Moroccan-Dutch speakers from Nijmegen (speaker cluster 5), realize a less outspoken length contrast between /ɑ/ and /a/, while making a very clear quality contrast between /ɑ/ and /a/ (one of the linguistic variables in cluster 2). In both respects, they differ from their Amsterdam counterparts (speaker clusters 3 and 4).

picture of Section 5.4.2.1, as no significant difference between the five speaker clusters was found. Variable cluster 3 is the one with the least pronounced social differentiation. Especially the duration of /ɑ/ and /a/ may be associated with some metrical (?) characteristic which has not been associated with any social evaluation or indexicality; alternatively, the functional load of the phonemic contrast is simply too high to allow for any socially meaningful variation.

From the number of pluses and minuses (in Table 6) for *variable cluster 4* it is clear that speaker cluster 1 has the largest mean value and this declines to the speakers in cluster 5, who have the lowest mean value. This picture is confirmed by the mean scores on the fourth variable cluster;[9] a Tukey Post Hoc analysis with variable cluster 4 as dependent variable reveals that speaker cluster 4 forms one subset with cluster 5 and at the same time another subset with cluster 3. So speaker cluster 4 is in between speaker clusters 3 and 5. With respect to variable cluster 4, speaker clusters 1 and 2 are separate subsets. The few T+M speakers in speaker cluster 2 (high-lighted in Figure 2) might have caused this speaker cluster's lower scores and thus it is not constituting a subset with speaker cluster 1. Another reason might be the related fact that for variable cluster 4 the Nijmegen C+D and T+M speakers (i.e. speaker clusters 2 and 5) have lower values than their respective Amsterdam counterparts.

So in variable clusters 1, 2 and 4, the role of linguistic coherence is scaled down by the differential social distribution in the use of the linguistic variables. Our findings reveal the multidimensionality in the linguistic and extra-linguistic patterns of co-variation; one cluster of linguistic variables (cluster 2) is areally defined, two others by the speakers' ethnic background in interaction with both their age and areal belonging (clusters 1 and 4).

For all clusters, the use of composite scores i.e. the sum of the linguistic variables, makes the social patterns, if present, more outspoken. This can be interpreted as a form of weak co-variation, in the sense that the composite scores disclose much more information than the linguistic variables in isolation. Forms of strong co-variation, holding within definable clusters of linguistic variables, occur wherever social and linguistic factors interact in their embedding and strengthen each other (van Hout, Kruijsen & Gerritsen 2014).

[9] Adding up the standardized scores for all linguistic variables belonging to the cluster gives a maximum 700 points for this variable cluster: in actual practice, speaker cluster 1 scores 654,19 > speaker cluster 2: 538,06 > speaker cluster 3: 451,70 > speaker cluster 4: 441,13 > speaker cluster 5: 365,18.

5.5 Conclusion and discussion

Witness the patterns uncovered in the data, the phenomena studied do not just constitute a feature pool in the sense of Mufwene (2001) or Cheshire et al. (2011), i.e. a pool of features with different roots in situations of language and dialect contact, where variation, from which speakers or speaker groups sample features to shape their social profile, is abundant. We also witness the formation of something Wiese (Wiese 2013:3) has described as a feature pond: "a pool that supports a network of interdependent features, a rich ecology that brings forth interconnected linguistic patterns at different levels."

In general and *a fortiori* in the study of emergent ethnolects the question arises: is it linguistics or rather ethnography that makes linguistic variables cluster? The two paradigms have existed side-by-side for decades (Labov and Trudgill versus Hymes and Gumperz, so to speak), but they have also taken root in relatively new research domains such as the study of ethnolects; whereas the ethnographic approach conceives language systems as infinite resources from which speakers may freely choose to shape their identity, the language centered approach tries to disentangle the rules, regularities and restrictions on these resources. Roughly, in the ethnographic approach, linguistically anything goes (or so it seems), whereas the language centered (or speech community-centered) approach primarily looks for the linguistic conditioning of the variation.

The first step in the analyses involved the cluster analysis of the 17 linguistic variables, based on 7 variable phenomena. As a statistical technique, based on similarities or dissimilarities between a set of variable phenomena, cluster analysis reveals patterns in grouping speakers and linguistic variables. We first focused on the linguistic variable clusters. The result, visualized in Figure 1 (Section 5.4.2.1), is without question interpretable and plausible from a linguistic point of view. The 17 linguistic variables constitute four clusters, the first three of which can unambiguously be labeled on linguistic grounds as the voicing of /z/ following inherently voiced segments (cluster 1), the PoA of monophthongal and diphthongal features (cluster 2) and vowel duration (cluster 3, which also contains the variable PoA of /a/, though). Linguistically, only cluster 4 is truly heterogenous, as it consists of morpho-syntactic and phonetic phenomena of several different types. With the exception of the use of *hun*, 'them', as subject, all variables have exogenous variants or are otherwise developing into ethnolectal features; the exogenous ethnolectal features originate in substrate effects (as e.g. the dentalization of /z/) or L2 acquisition (the marking of grammatical gender).

We can thus conclude that on this overall level both linguistic and social or

ethnographic factors make the linguistic variables cohere, although the linguistic rhyme and reason is the first one to catch the eye – since linguistically three of four clusters of variables are fairly homogenous.

The second step in the analyses involved (i) the cluster analysis of the 51 speakers in the sample (see Figure 2 in Section 5.4.2.2) and (ii) relating the resulting five speaker clusters with the four linguistic variable clusters (Table 6 in Section 5.4.3).[10] The findings allow us to determine to which extent the clustering of the linguistic variables (resulting from the first step) applies to each of the speaker clusters.

It appeared that none of the five clusters of speakers treats the linguistic variables in all four clusters consistently. For linguistic variable cluster 2, which is largely rooted in the distinction between the two urban dialects involved, the Turkish-Dutch and Moroccan-Dutch speakers from Nijmegen and Amsterdam resemble their respective local white Dutch peers. For the linguistic variable cluster 3, there are relatively few (and definitely no systematic) differences between the five speaker clusters; all speaker clusters use at least 75% standard variants for three of the linguistic variables in linguistic variable cluster 3 (it is impossible to identify the standard variant for /a/ PoA, the fourth variable in this cluster; see Section 5.3.4).

Only variable clusters 1 and 4 show pronounced ethnic differentiations. For all seven linguistic variables in cluster 4, the white Dutch speakers are clearly on the standard side of the scales, whereas the Turkish-Dutch and Moroccan-Dutch speakers are so in about half of the cases. This social distribution is probably related to the fact that, with one exception, all phenomena in this variable cluster have exogenous variants, originating in substrate effects or L2 acquisition or are otherwise developing into ethnolectal features. The endogenous exception, the use of *hun* as a 3 pl. subject pronoun, is a recent change in colloquial spoken Dutch, strongly opposed by teachers and the intellectual elite. It is presently gaining ground in younger generations and its use seems to be incremented (Labov 2001) by non-white groups.

Remarkably, the 12 year old Amsterdam Turkish-Dutch and Moroccan-Dutch speakers and all Turkish-Dutch and Moroccan-Dutch speakers from Nijmegen are on

[10] In this regard, our study resembles the remarkable and highly advanced analysis presented by Ma & Herasimchuk (1972) of 45 bilingual Puerto Rican speakers of English and Spanish in New York City. On the basis of pooled production data for a range of variable phenomena in both languages, the authors ran factor analyses for the linguistic variables and Q analyses for the speakers. Cross tabulating the resulting clusterings for the cases that were represented by substantial amounts of data, the authors ultimately drew up the profiles in the use of five clusters of linguistic variables of four different speaker groups (p. 285). Interestingly three of the five clusters of linguistic variables consist of both English and Spanish variable phenomena.

the standard side of the scales in linguistic variables cluster 1, whereas the 20 year old Amsterdam Turkish-Dutch and Moroccan-Dutch speakers as well as all white Dutch speakers are on the far non-standard side of the scales. The linguistic variables in cluster 1 concern the voicing of /z/ after inherently voiced segments and in this connection the behavior of the white Dutch speakers is in line with the general trend (in a huge part of the Dutch language area) to devoice fricatives across the board, a change in the direction of the typologically and phonetically unmarked case.[11] Since the older standard is subject to ongoing change (from below) in the white Dutch community but is being retained by the immigrants, what was the norm is now used mostly by immigrants, while the white Dutch have moved on in a new change, a clear example of the issue of (non-) diffusion of a change across certain ethnic groups. Our findings suggest that in this respect the traditional standard norm is in the process of becoming 'ethnolecticized', i.e. subject to a process of social redistribution, the outcome of which might be that the phenomenon at hand, the voiced realization of /z/ after inherently voiced segments, is the exclusive characteristic of specific ethnic groups.

In short, on the overall level the coherence of linguistic variables in the varieties studied appears to be linguistically determined in the first place. Closer scrutiny shows that for a small part of the picture, there are hardly any systematic differences between the various groups of speakers. For another part, 'ethnolectization' is occurring, either as a result of an emerging 'superdiversity' in microcosmos, resulting from the addition of exogenous variants to prior-existing linguistic variables, here the dentalization of /z/ (cf. Hinskens 2013), or through the ethnic redistribution of the social stratification of endogenous standard (voicing of /z/) or non-standard variants (monophthongization of /ɛi/). These developments may herald a new, partly ethnically-bound cohesion among linguistic variables.

[11] Fricative noise requires a high volume of airflow to create the turbulence that is perceived as the fricative quality, but voicing impedes the airflow and weakens the fricative noise component. So in a sense voiced fricatives are acoustically self-contradictory.

6 Conclusion and discussion

6.1 Recapitulation

In Chapter 1 the concept of ethnolects was explored, and an overview of the ethnolects of Dutch as well as the state of affairs in international research was given. This was followed by a description of the overarching project 'Roots of Ethnolects' with youngsters from three backgrounds ('white' Dutch, Turkish-Dutch and Moroccan-Dutch), two cities (Amsterdam and Nijmegen), and two age-groups (10-12 and 18-20 year olds). The 'core corpus' was introduced containing a selection of speakers from the Roots of Ethnolects spontaneous speech database. In the next subsections, a recapitulation is given of the results of Chapter 2 (/z/), Chapter 3 (/ɛi/), Chapter 4 (/a:/ and /ɑ/) and Chapter 5 (co-variation).

6.1.1 /z/

Chapter 2 examined the voiced fricative /z/ in word-initial position. We first investigated the occurrence of a non-standard Dutch dentalized /z/ and subsequently we established whether Dutch processes of devoicing were applied.

We found that the Moroccan-Dutch and Turkish-Dutch youngsters employ dentalization, but the 'white' Dutch do not do so. Dentalization in the realization of /z/ appears to depend on the background of the speaker as well as the background of the interlocutor. The Moroccan-Dutch and Turkish-Dutch speakers use the least amount of dentalization when speaking to 'white' Dutch interlocutors and the most when speaking to Moroccan-Dutch interlocutors, which is in line with the amount of dentalization of the three groups of speakers: 'white' Dutch < Turkish-Dutch < Moroccan-Dutch. The behavior of the Moroccan-Dutch is most pronounced as they used more dentalized realizations than the Turkish and they used them more in conversations with Moroccan-Dutch interlocutors. For the Turkish-Dutch, the dentalization patterns differ for both cities. In Amsterdam, the pattern of the Turkish-Dutch speakers with regard to interlocutors is Turkish-Dutch < Moroccan-Dutch / 'white' Dutch, in Nijmegen the pattern is Moroccan-Dutch / 'white' Dutch < Turkish-Dutch. The differential patterning of dentalization for both groups suggests that it is a relatively stable and probably older in-group feature for the Moroccan-Dutch, the difference between the Amsterdam and Nijmegen Turkish-Dutch suggests that it is a more recently acquired identity-marking resource for the Turkish-Dutch.

Devoicing is dependent on the linguistic context (post-obstruent vs. post-sonorant and post-vowel). The Moroccan-Dutch and Turkish-Dutch speakers do use voiced variants of /z/ in the post-obstruents context, while the 'white' Dutch hardly do so, as expected. Just as in the post-obstruent context, the 'white' Dutch mostly used

devoiced variants in the post-sonorant and post-vowel context, while the Turkish-Dutch mostly used voiced ones. In the post-sonorant + post-vowel context, Amsterdam speakers showed more devoicing than the Nijmegen speakers. This is in particular the case for the 18-20 year olds.

6.1.2 /ɛi/

Chapter 3 examined the Dutch diphthong /ɛi/. Two variable properties of /ɛi/ were examined: (1) the degree of monophthongization, and (2) height of the prominent, first element.

The most important result for monophthongization of /ɛi/ is the difference between the two 'white' Dutch groups (with and without ethnic ties) on the one hand and the Turkish- and Moroccan-Dutch speakers on the other hand, with the latter two groups showing more monophthongization. There was also a moderate difference between the 20 year olds and the 12 year olds, with the 20 year olds using more monophthongized variants. Monophthongization is sensitive to etymological distinction between <ei> and <ij>, the position vis-à-vis the right syllable boundary (i.e. open vs. closed syllables) and the frequency of occurrence of the citation forms. However, the effects of these linguistic factors seem to run parallel to the intricate effects related to the background of the interlocutor.

The most outspoken extra-linguistic effect for the height of the prominent first element of /ɛi/ is that of city. For all groups, we find that the Amsterdam speakers use more open first, prominent elements of the diphthong than the Nijmegen speakers. The 20 year old speakers use more open prominent, first elements of the diphthong than the 12 year olds.

There are also complex style effects in the realization of the diphthong /ɛi/ which are connected with the background of the interlocutor. For the monophthongization of /ɛi/ this style-as-accommodation effect is found only in the speech of the 20 year old Turkish-Dutch speakers from Amsterdam. For lowering, only the 20 year old Turkish-Dutch cohort from Nijmegen showed this style effect.

There is a very clear tendency towards a double resetting of the social distribution of the variation in the realization of the diphthong in local dialect variants; they appear to be changing from sociolectal into ethnic markers.

6.1.3 /a:/ and /ɑ/

Chapter 4 examined two Dutch phonemes, /a:/ and /ɑ/, and their distances in two linguistic conditions, before obstruents and before nasals, and two dimensions, duration and place of articulation.

City has an important influence on all four place of articulation variables (PoA /ɑ/

obs, PoA /ɑ/ nas, PoA /a:/ obs, PoA /a:/ nas). In the case of place of articulation of /ɑ/, Amsterdam speakers had a lower mean value than the speakers from Nijmegen, while for /a:/ it is the opposite. Amsterdam speakers also had a higher value on duration of /ɑ/ before nasals.

Age had an effect on duration of /ɑ/ before nasals, the 12 year olds using on average longer /ɑ/s before nasals than the 20 year olds, and on place of articulation of /a:/ before obstruents, the 20 year olds having a higher value than the 12 year olds.

For the Moroccan, Turkish and 'white' Dutch with inter-ethnic ties, no main effects of the background of the speaker or the interlocutor was found for both phonemes, neither in the two linguistic conditions nor for the two dimensions. However, both background of the speaker and of the interlocutor were involved in interaction effects. With regard to the two 'white' Dutch groups, a difference was found between the two groups for place of articulation of /a:/ before obstruents. The speakers with inter-ethnic ties had a higher mean in in-group conversations than the speakers without inter-ethnic ties.

No effects were found for the distance between the durations of /ɑ/ and /a:/. For the distance between the place of articulation of /ɑ/ and /a:/, effects of city and background of the speaker were found, as well as an interaction between those two. Amsterdam speakers differentiated less between the two vowels than the Nijmegen speakers. However, the 'white' Dutch speakers without inter-ethnic ties ('C-group') of Amsterdam differed from the Turkish and Moroccan Amsterdam speakers and in case of the obstruent condition also from the 'white' Dutch speakers with inter-ethnic ties ('D-group') of Amsterdam. In fact, they seem to pattern with the Nijmegen speakers instead of with their own Amsterdam peers. The distances between the place of articulation of /ɑ/ and /a:/ are smaller for the 20 year old speakers than for the 12 year olds.

6.1.4 Co-variation

Chapter 5 examined the co-variation between 17 linguistic variables. These 17 linguistic variables were found to group together into four variable clusters. *Variable cluster 1* contains two phonetic variables concerning the voicing of /z/ following inherently voiced segments. *Variable cluster 2* contains four variables that all involve phonetic variation in connection with the place of articulation (high/low, front/back) of the relevant vowels. These four variables are also the variables for which the most important extra-linguistic factor is city. In *variable cluster 3* all but one linguistic variables have something to do with the duration of vowels. The three variables that involve duration do not show any social effect. *Variable cluster 4* contains all four

morpho-syntactic phenomena (all in the domain of inflection) as well as three phonetic phenomena which are all variables with ethnolectal variants. For these variables, speaker background is the most important extra-linguistic factor.

On the basis of the behavior of the speakers with regard to the 17 linguistic variables, five speaker clusters can be distinguished. A clear division was found between the 'white' Dutch speakers (*speaker cluster 1 and 2*) and the Turkish-Dutch and Moroccan-Dutch speakers (*speaker clusters 3, 4 and 5*). *Speaker cluster 1* (labelled 'C+D, A+N') contains all 'white' Dutch speakers from Amsterdam as well as half of the 'white' Dutch speakers from Nijmegen. The other half of the 'white' Dutch speakers from Nijmegen are part of *speaker cluster 2* ('C+D, N'). *Speaker clusters 3 and 4* both contain Turkish-Dutch and Moroccan-Dutch speakers from Amsterdam, *cluster 3* ('T+M, A20') mainly 20 year olds, *cluster 4* ('T+M, A12') mainly 12 year olds. *Speaker cluster 5* ('T+M, N') contains Turkish-Dutch and Moroccan-Dutch speakers from Nijmegen.

Section 5.4.3 answers the question how the clusters of linguistic variables and the clusters of speakers are related. Our findings reveal the multidimensionality in the linguistic and extra-linguistic patterns of co-variation; one cluster of linguistic variables (variable cluster 2) appears to be areally (the two cities) defined, two others by the speakers' ethnic background in interaction with both their age and areal belonging (variable clusters 1 and 4).

6.2 How different are the groups: co-variation again

Is it possible to get a better view on the differences between the two cities and the groups? Chapter 5 presented cluster analyses to show how the linguistic variables covary and how groups of speakers can be distinguished. Here we present an additional statistical analysis that yields a more refined picture of the groupings of the speakers involved in this study. It has been visualized in Figure 1 below.

Figure 1 displays the speakers in a two-dimensional space. Each point represents one of the 51 speakers. The closer speakers are diagramed, the more alike they are. Each speaker is marked for a 'white' Dutch (round point) or Turkish-/Moroccan-Dutch background (square point) and for city (i.e. grey for Amsterdam, white for Nijmegen). The result makes roughly a fourfold division of the speakers based on the city and background of the speakers. Clockwise, starting at the top left-hand corner, we notice a group of Turkish-Dutch and Moroccan-Dutch speakers from Nijmegen. Next, at the top right-hand corner, there is a group of Turkish-Dutch and Moroccan-Dutch speakers from Amsterdam. Below, at the bottom right-hand, there is a group of first and foremost 'white' Dutch speakers from Amsterdam. The bottom left-hand group

also contains predominantly 'white' Dutch speakers, but from Nijmegen.

This fourfold division is also visible when we divide the speakers based on their values on variable clusters 2 and 4. As we saw in Section 5.2.1, Table 4, city had a strong effect on variable cluster 2, while the background of the speaker had a strong effect on variable cluster 4. When looking at the dividing line of variable cluster 2 in Figure 1, it is noticeable that all but one Nijmegen speakers are on one side of the dotted line, while all but one Amsterdam speakers are on the other side. Recall that variable cluster 2 involved phonetic variation in connection with the place of articulation (high/low, front/back) of the vowels (cf. Section 5.2.1) that is largely rooted in the distinction of the two urban dialects involved. With regard to the dividing line of variable cluster 4 in Figure 1, most Turkish and Moroccan speakers are on the top side of the dashed line, while 23 of the 25 'white' Dutch speakers are on the bottom side. Variable cluster 4 showed pronounced ethnic differentiations.

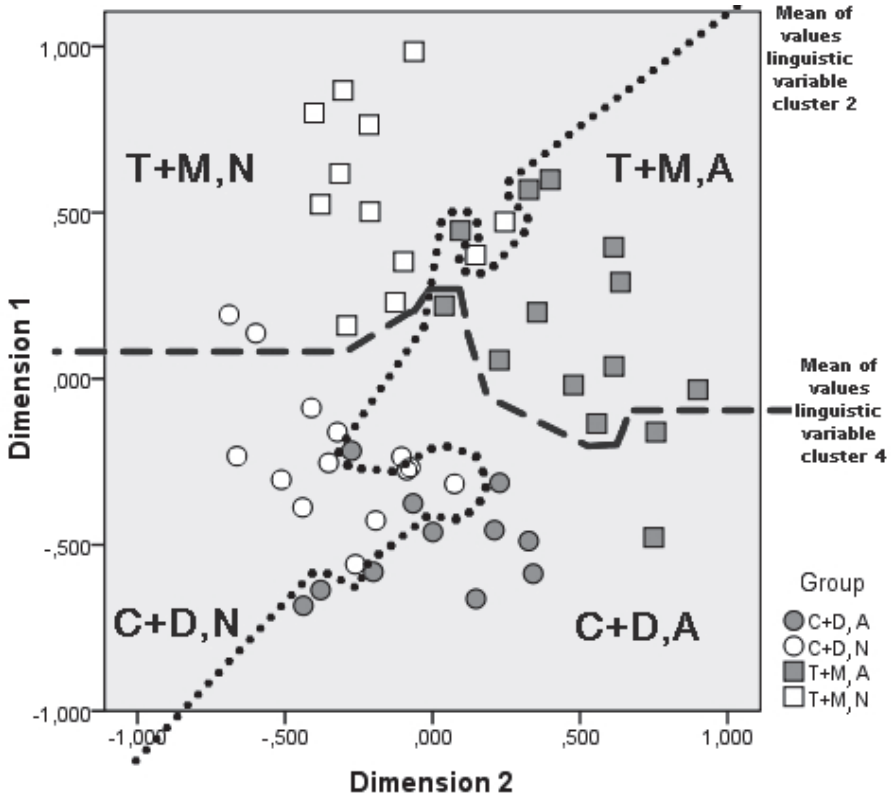


Figure 1 Plot of the 51 speakers on the first two dimensions of MDS analysis, with two inserted division lines based on the speakers' values on variable clusters 2 and 4 (see Chapter 5). C = 'white' Dutch without inter-ethnic ties, D = 'white' Dutch with inter-ethnic ties, T = Turkish-Dutch, M = Moroccan-Dutch, A = Amsterdam, N = Nijmegen

These two divisions can be seen as the result of an old, endogenous Dutch dialect-geographical division, the characteristics of which return in the linguistic behavior of the youngsters we investigated. This dialectal distinction is crossed by a much younger ethnolectal division that connects the ethnic groups over and above the dialectal distinctions, although the urban areas of Nijmegen and Amsterdam are geographically clearly separated.

6.3 The research questions

For the study in this thesis, we focused on seven research questions. Four of them addressed the roots, i.e. the origin of the linguistic variation patterns. The fifth research question focused on style shifting in connection with the background of the interlocutor and the last two deal with the social diffusion of patterns of variation. Implications of the outcomes for each of the seven research questions are discussed below, in relation to each of the seven research questions.

Roots / Origins

1. Substrate effects

To what extent can variation patterns be related to interference from the heritage languages of speakers with Moroccan and Turkish backgrounds?

Both the voicing of /z/ after obstruents and the dental realization of /z/ are exogenous, 'un-Dutch' phenomena. Post-obstruent voicing of /z/ definitely does not have Dutch roots. The dental realization of /z/ has its roots in the languages spoken by the Moroccan-Dutch and has been borrowed by the Turkish-Dutch.

2. Regional effects

To what extent can variation patterns be related to the impact of the local nonstandard varieties, i.e. to the dialect differences between the cities of Nijmegen and Amsterdam?

Amsterdam speakers showed more devoicing of /z/ in the post-sonorant + post-vowel context than the Nijmegen speakers, especially in the case of the 18-20 year olds. Also, the Amsterdam speakers used more open first, prominent elements of the diphthong /ɛi/ than the Nijmegen speakers. They also have longer /ɑ/s before nasals, more fronted /ɑ/'s (in both obstruent and nasal context) and more backed /a:/s in closed syllables (in both contexts) than the Nijmegen speakers. This is in line with the local nonstandard varieties.

Nijmegen speakers have a more open pronunciation of /ɛ/ in the phrases *heb je/ik* 'have you/I'. Together with three other variables (/ɑ/ PoA, /ɛi/ height of first element, and difference between PoA /ɑ/ and /a:/) that involve phonetic variation in

connection with the place of articulation (high/low, front/back) of the relevant vowels, the height of /ɛ/ forms a cluster of linguistic variables that have different pronunciations in the Amsterdam and Nijmegen sociolects. There was a clear division line between the Amsterdam speakers and Nijmegen speakers in our data for this cluster.

3. Structural effects

To what extent does endogenous Dutch linguistic conditioning apply in the patterns of variation in speakers with Moroccan and Turkish backgrounds? Are conditions reset?

This question is particular relevant to the devoicing of /z/. According to endogenous Dutch linguistic conditioning (on the surface, obstruent clusters containing one of more fricatives are categorically voiceless), /z/ should be devoiced after obstruents. The devoicing of /z/ following obstruents occurs less frequently among the Turkish-Dutch and Moroccan-Dutch speakers.

With regard to the diphthong /ɛi/, monophthongization turns out to be sensitive to the position vis-à-vis the right syllable boundary, the etymological distinction between <ei> and <ij> (which is still phonological in several groups of Dutch dialects) and the frequency of occurrence of the citation forms. However, if one looks in more detail, it is hard to interpret the results. The corpus of speech data under study is unfortunately too small to investigate these interesting linguistic conditioning in more detail.

4. Second language acquisition effects

To what extent can variation patterns be related to properties resulting from processes of second language acquisition?

This question is relevant to the distances between /a:/ and /ɑ/ as we expected an influence from second language acquisition, induced by the typological distinction between syllable- versus stress-timed systems between Dutch on the one hand and the heritage languages on the other; unlike Dutch, the heritage languages do not have any phonological contrast between lax and tense vowels. However, the Turkish-Dutch and Moroccan-Dutch speakers happen to make the same length difference as their ‘white’ peers. So, the expected L2 effect did not manifest itself.

Turkish- and Moroccan-Dutch speakers use more non-standard Dutch *hun* as a 3 plural subject pronoun than the two ‘white’ Dutch groups. This use as subject pronoun is a recent change in colloquial spoken Dutch and its use seems to be incremented by non-white groups, but that is not a second language acquisition effect.

Position

5. Style-shifting effects

To what extent do speakers shift between patterns of variation, depending on the background of the interlocutor?

The background of the interlocutors had several influences on the linguistic behavior of the speakers. It was most pronounced for the dentalization of /z/. Speakers from the Turkish-Dutch and Moroccan-Dutch groups showed the lowest dentalization indexes in their conversations with 'white' Dutch and the highest indexes in conversations with speakers of the same background. This behavior is even more pronounced for the Moroccan-Dutch than that for the Turkish-Dutch. Dentalization seems to be an in-group feature for both groups. Compare Fagyal & Stewart's (2011:75) finding of an intonational variant which "seems to function as a micro-level style feature indexing common ground" among preadolescents in a Paris *banlieu*. There are complex style effects in the variation in the realization of the diphthong /ɛi/ which are connected with the background of the interlocutor. The patterns are sometimes hard to interpret because there tend to be several effects at the same time and complex interaction effects. Also for the realization of the vowels /ɑ/ and /a:/, several style effects were found that are related to the background of the interlocutor, for both duration and place of articulation and in both condition (before obstruent and before nasal).

Bell's (1984) insight that characteristics of the speech used in interaction with members of certain groups tends to mirror characteristics of the speech of members of these groups themselves, is most clearly reflected in the dentalization of /z/. Speakers use increasing amounts of dentalized /z/ when speaking to members of the three groups in line with the pattern "'white' Dutch < Turkish-Dutch < Moroccan-Dutch", which mirrors the degrees of dentalization in the speech of the members of these three groups. As for the other vowels variables, Bell's (1984) insight applies less clearly. It showed up only in the variation in the monophthongization of /ɛi/ by the 20 year old Turkish-Dutch speakers in Amsterdam, and in the variation in the duration of /a:/ before obstruents by the Moroccan-Dutch. In the latter case, the speech of the Moroccan-Dutch towards their interlocutors replicate the pattern of the variation in the duration of /a:/ in the speech of the members of the respective groups: Moroccan Dutch ≤ Turkish Dutch < white Dutch.

Social spread

6. Inter-ethnic convergence effects

Do speakers with Moroccan and Turkish backgrounds differ in their variation patterns or do they share a more global 'non-native' identity?

There is hardly any difference between the Turkish-Dutch and the Moroccan-Dutch speakers on the phonemes investigated. It looks as if dentalization of /z/, which has Moroccan Arabic and Berber roots, has come to serve as a marker of a more global 'non-native' identity and has hence started to develop into a 'multi-ethnolect' feature. However, the Moroccan-Dutch are leading the use of dentalized variants, as they are more able to adapt dentalization depending on the interlocutor than are the Turkish Dutch. The monophthongization of /ɛi/ seems to be ethnolectized by the Turkish-Dutch and Moroccan-Dutch speakers, while the monophthongal variant is being rejected by 'white' Dutch speakers, i.e. it is going through a social redistribution.

7. Native convergence effects

Is there any evidence of spread of ethnic patterns of variation to 'white' Dutch peers?

The 'white' Dutch speakers do not adopt the ethnic dentalized variants at all. These Dutch speakers also seem to keep to their own track in the voicing of /z/ as they devoice more than the Turkish-Dutch and Moroccan-Dutch speakers.

As said before, monophthongization of /ɛi/ is going through a social redistribution by which the monophthongal variant is rejected by the 'white' Dutch speakers and ethnolectized by the Turkish-Dutch and Moroccan-Dutch speakers. So for the phenomena under study no 'crossing' (Rampton 1995) seems to occur. There is a clear dividing line between the 'white' Dutch on one hand and the Turkish- and Moroccan-Dutch on the other.

6.4 Discussion

The Roots of Ethnolects project has a structural-linguistic approach to investigate modern Dutch ethnolects and the origin of their features. This has not been done earlier in the Dutch language area. The studies in the previous chapters provide new insights in the systematic nature of ethnolectal varieties, both from a linguistic and social point of view, especially with respect to sociophonological aspects.

The focus was on variable features that can be supposed to be 'below the level of conscious awareness' (or which were at most semi-conscious) for speakers. Therefore words that were uttered in an (intentionally) conspicuous way (for the most part imitations) were not selected and neither were words which were read from, for example, papers and magazines. In this way, we tried to focus on features of what is

called *primary* ethnolect by Auer (2003), while excluding possible features of 'mediatized' and hence stylized varieties, labeled *secondary* ethnolect by Auer (2003), as well as excluding possible features of a *tertiary* ethnolect. The latter might be the type of *wannabe* ethnolect (cf. "Murks", Nortier 2002) which could be used by native youngsters without immigrant backgrounds who have no direct contact with speakers of the *primary* ethnolect, i.e. e.g. the Dutch speakers without inter-ethnic ties ('C-group'). While it might be very interesting to see if our Dutch speakers without inter-ethnic ties use any *tertiary* ethnolect, i.e. supposedly ethnic features picked up from the media, the speakers in our research merely function as a control group for the Dutch with inter-ethnic ties, to establish whether and how the latter accommodate to Turkish-Dutch and Moroccan-Dutch speakers. Their accommodation to ethnolectal varieties appears to be only marginal. The accommodation is rather the other way around. Turkish-Dutch and Moroccan-Dutch speakers adopt many features from the surrounding urban dialects. They become the co-carriers of accent features which have a dialect-geographical origin and at least one which (the monophthongization of /ɛi/) has developed into a sociolect feature.

In this thesis the variation in the realization of the phonemes /z/, /ɛi/, /a:/ and /ɑ/, as well as the contrast between the latter two was examined. All four segments are fairly to highly frequent in (colloquial) Dutch speech. The phonemes were chosen to represent three different kind of roots (i.e. origins). The dentalized variant of /z/ has roots in the heritage languages of the Turkish- and Moroccan-Dutch speakers. The variation in the realization of /ɛi/ is rooted in local urban dialects, while changes in the contrast between /a:/ and /ɑ/ is supposed to be rooted in second language acquisition. The expected effect of L2 acquisition on the /a:-ɑ/ contrast was not found, which might be a disadvantage of the chosen approach[1] as only one variable pair was examined that may be affected by L2 acquisition processes. It seems worthwhile nevertheless to include other tense-lax pairs as modern Dutch is involved in a chain shift involving the diphthongization of the lax long mid-vowels and the lowering of the (homorganic) diphthongs. Diphthongization seems to take over the role of length and it would be interesting to know how ethnolect speakers cope with these changes.

Another topic to be tackled is the apparently leading role that Moroccan-Dutch speakers have in shaping a multi-ethnolect variety. Amsterdam has relatively more inhabitants with a Moroccan background, Nijmegen more inhabitants with a Turkish

[1] However, in another study within our project, an L2 acquisition effect was found in the variation in the marking of Dutch grammatical gender, which involves a range of related morpho-syntactic phenomena (cf. Hinskens et al. in preparation).

background. Nevertheless, the impact of Moroccan-Dutch speakers seems to be higher in both urban areas. This may be caused by the stronger position Dutch has as a lingua franca between speakers with a Moroccan background. For them Dutch is more important both as a means of communication (Morocco is a multilingual country) and as a vehicle of their identity.

The difference between the two groups is confirmed by the findings in Grondelaers, van Gent & van Hout (2015). They took 72 speech clips from 18-20 years old speakers from the Roots of Ethnolect database and presented them to university students (n=21) with the question to determine the ethnic background of the speakers of the clips (Moroccan, Turkish or Dutch) and the city of origin (Amsterdam or Nijmegen). The 'white' Dutch speakers were perfectly recognized by the students. Some Turkish-Dutch speakers were classified as ‘white’ Dutch too, which hardly happened to the Moroccan-Dutch speakers. The choice for Moroccan-Dutch speakers was more often Moroccan, whereas the choice was in balance for the Turkish speakers. Moroccan seems to be a more attractive label for ethnolectal speech, though the distinction between Turkish and Moroccan is hard to recognize. The recognition of the urban area was successful. This implies that both the ethnolectal and dialect-geographical division lines (as visible in Figure 5 of Chapter 3 and Figure 1 in Chapter 6) were reproduced in the perception study of Grondelaers et al. (2015).

The finding that the Moroccan- and Turkish-Dutch speakers are recognized as Dutchmen of Moroccan and Turkish extraction and as inhabitants of Amsterdam or Nijmegen is supported by the inventory drawn up by Geval (2011) of answers to questions related to Moroccan- and Turkish-Dutch in a written survey. One in six respondents, mostly dialect speakers, were able to answer the question: "Are there people in your area with a Turkish or Moroccan background of whom you can hear that they come from your hometown or region?” and to give some examples. Most examples concern the pronunciation of certain speech sounds. The respondents were also asked what is striking to them about the speech of Moroccan- and Turkish-Dutch people. Over 50% of the respondents was able to answer this question and came up with no fewer than 231 different characteristics. The most common ones (mentioned 113 times) are related to the pronunciation of certain sounds.

Finally, some attention should be paid to the database. In the ideal situation, we would have a database with 96 speakers (cf. Table 1 of Section 1.4.2) of which all speakers with interethnic ties (the D-, T- and M-groups) would have had conversations with interlocutors from all three backgrounds and all speakers without inter-ethnic ties (C-group) with two of their own peers. As was pointed out in Section

1.4.3, this ideal situation has not been reached. Given this fact and the fact that, due to the amount of available transcribers and time limits, controlled transcriptions of at least 10 minutes of dialogical, relatively spontaneous natural speech were not available for all conversations, we opted for the selection of a 'core corpus'. A complicating factor is that the groups (defined by the extra-linguistic variables city, background and age group) were quite small with about three speakers per group (cf. Table 4 of Section 1.6) and this has consequences for the robustness of certain effects. Several style effects in connection with the interlocutors' backgrounds were found, but because of the small groups and the statistical methods used, it was sometimes hard to make generalizable substantial claims about these effects, as was for example the case for the monophthongization of /ɛi/.

6.5 Further research

In Chapters 2, 3 and 4 we already gave some specific suggestions for further research. In short, Chapter 2 asked for further research of voiced /z/ in Turkish compared to the other languages; Chapter 3 suggests to broaden the empirical basis of the study of /ɛi/ by including more speakers for each of the relevant groups defined by the variables city, background and age group; Chapter 4 proposes follow-up research using the 'Roots of Ethnolects' data: study of other tense/lax vowel contrasts in Dutch, such as /ɔ/ versus /o:/ and /ɪ/ versus /e:/, and other potentially variable features defined by the typological distinction between stress-timed and syllable-timed systems.

There are also larger, more general questions that relate to future research. The overarching project 'Roots of Ethnolects' focused on more questions than the six addressed in our studies. Section 1.41 gives al nine questions, three of which we did not address in this thesis.

The first one concerns the aspects of language use (components of the grammar) that characterize ethnolects as distinct varieties. We focused on sociophonological aspects, as well as some morphosyntactic variables in the covariation study. Muysken (2013) and Hinskens (2011) gave overviews of the various phenomena encountered in the recordings in the Roots of Ethnolects database that give a first impression to answer this question. Most of these phenomena (in the sound components or in morpho-syntax) still need to be further and systematically examined. This seems to apply in particular to prosodic aspects, although syntax is an interesting area of research in ethnolectal studies as well. Although Kiezdeutsch (cf. Wiese 2009) contains a number of deviant word order patterns and new construction types, deviant syntactic patterns, of the type that often originates in second language acquisition, hardly seem to occur in our recorded

samples of Moroccan and Turkish Dutch.

The second question we did not address is the distribution of ethnolects over the Netherlands. The project focused on Amsterdam and Nijmegen, but it is obvious that more Dutch cities should be included in future research. Such research is needed to find out how ethnic and geographical boundaries interact and maybe even strengthen each other. On a higher level of abstraction it would probably be revealing to compare our findings to similar urban areas in Germany, England, Sweden, etc. Unmistakably, international research cooperation (maybe starting at the level of the Germanic language group) could greatly help to unravel internal, external and extra-linguistic forces underlying the use and development of new nonstandard features as well as their diffusion to the verbal repertoires of other cultural groups. (Cf. e.g. Nortier & Svendsen (eds.) 2015, which allows the comparison of outcomes of ethnographic studies of youth languages spoken around the globe, mainly in Europe and Africa.)

The third question left concerns the extent to which ethnolects resemble code-switched varieties and contain overt lexical and grammatical material from several different languages. It seems that in the in-group conversational data, much more code-switching and code-mixing occurs among the Turkish-Dutch than among the Moroccan-Dutch. However, this impression is not yet quantitatively underbuilt and needs to be investigated.

As became clear from the demographical description in Section 1.5.1, more people with a foreign origin live in Amsterdam (around 51%) than in Nijmegen (around 25%). Also (as mentioned above), the proportions of inhabitants with a Moroccan origin and those with a Turkish origin is reversed in both cities, with more Moroccans than Turks in Amsterdam and more Turks than Moroccans in Nijmegen. It is not clear if (or how) this had any effect on the speech of our speakers. In Chapter 2, a difference was found for the Turkish-Dutch speakers in Amsterdam and those in Nijmegen. The Turkish-Dutch speakers in Amsterdam used more dental /z/ when speaking to Moroccan-Dutch interlocutors than to the other two groups, while the Turkish-Dutch speakers in Nijmegen use most dental /z/ when speaking to fellow-Turkish Dutchmen of the same age. Dental /z/ seems to function as a *trait d'union* with the Moroccan-Dutch in Amsterdam, while it seems to function as an in-group factor in Nijmegen. So, with regard to dental /z/, the Turkish-Dutch speakers are more oriented towards the Moroccan-Dutch group in a city were more Moroccan-Dutch than Turkish-Dutch live, while being more oriented towards Turkish-Dutch in a city were more Turkish-Dutch than Moroccan-Dutch live. This rather speculative explanation obviously needs to be tested in further research, preferably also on other variable phenomena.

Another important question for the future is the role of perception. As we saw in Chapter 2, the dental /z/ functions as an ethnic marker for the Turkish-Dutch and Moroccan-Dutch speakers. It would be interesting to see if this is a typical characteristic that is perceived and recognized by speakers of several backgrounds (whether 'white' Dutch, Turkish-Dutch, Moroccan-Dutch, or even other backgrounds). For example, if exposed to speech of Turkish-Dutch and Moroccan-Dutch speakers, would any speaker of modern Dutch mention the (dental) /z/ as sounding e.g. 'non-Dutch' or 'Moroccan'? And which other elements are characterized as such? From Geval's (2011) study it appears to be sufficiently salient to be explicitly mentioned by several 'white' Dutch dialect speakers.

Grondelaers et al. (2015) had people rate fragments of Moroccan Dutch speech on a number of evaluation scales. It turned out that the Moroccan-Dutch speakers (in the Roots of Ethnolect database) were systematically downgraded on status. They are evaluated by 'white' Dutch speakers as having low status. The question is how they are evaluated by ethnolect speakers and how strong their covert prestige is. This relates to the question of the larger landscape of Dutch varieties. How do ethnolect varieties fit in the landscape that will evolve in the near future?

Finally, we wish to point out a very concrete matter. As transcribing is laborious and time-consuming, some conversations of speakers who do not belong to our 'core corpus' still need to be transcribed, or their transcriptions need to be extended to at least 10 minutes; in some cases, transcriptions of 10 minutes of speech or more are only waiting for control by a second transcriber. Once it has been established that all conversations contain at least 10 minutes of controlled transcriptions, they can be used for further research, either to extend the current study or to examine other features that have not yet been studied.

As was mentioned in the previous section, not all of the speakers in the Roots of Ethnolects spontaneous speech database (cf. Table 2 of Section 1.4.3) participated in conversations in all three contact situations. While the data of these speakers can therefore not be taken into account for the further study of potential style effects caused by the interlocutor, they are relevant for any study on ethnolectal variation that does not involve the background of the interlocutor as a factor. In short, it is definitely worthwhile to further extend the data for the Roots of Ethnolects database on the basis of the available recordings.

References

Abdel-Massih, Ernest. 1973. *An introduction to Moroccan Arabic*. Michigan: Center for Near Eastern and North African Studies, University of Michigan.

Abercrombie, David. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Androutsopoulos, Jannis. 2001. From the streets to the screens and back again: On the mediated diffusion of variation patterns in contemporary German. Essen: LAUD Linguistic Agency/Universität Essen.

Auer, Peter. 1997. Co-Occurrence Restrictions between Linguistic Variables: A Case for Social Dialectology, Phonological Theory and Variation Studies. In Frans Hinskens, Roeland van Hout & W. Leo Wetzels (eds.), *Variation, Change, and Phonological Theory*, 69–99. (Current Issues in Linguistic Theory 146). Amsterdam: John Benjamins Publishing Company.

Auer, Peter. 2003. "Türkenslang": ein jugendsprachlicher Ethnolekt des Deutschen und seine Transformationen ["Türkenslang": a youth language ethnolect of German and its transformations]. In Annelies Häcki Buhofer (ed.), *Spracherwerb und Lebensalter*, 255–264. Tübingen/Basel: Francke. http://www.forum-interkultur.net/Beitraege.45.0.html?&tx_textdb_pi1[showUid]=8 (29 September, 2010).

Baker, Philip. 2000. Theories of creolization and the degree and nature of restructuring. In Ingrid Neumann-Holzschuh & Edgar Schneider (eds.), *Degrees of restructuring in creole languages*, 41–63. Amsterdam/Philadelphia: Benjamins.

Becker, Kara. 2015. Linking community coherence, individual coherence, and bricolage: The co-occurrence of (r), raised BOUGHT and raised BAD in New York City English. Lingua [Special Issue: *Frans Hinskens & Gregory R. Guy (eds.),* Coherence, covariation and bricolage. Various approaches to the systematicity of language variation]. doi: 10.1016/j.lingua.2015.10.017

Bell, Allan. 1984. Language Style as Audience Design. *Language in Society* 13(02). 145–204. doi:10.1017/S004740450001037X.

Bellmann, Günter. 1998. Between Base Dialect and Standard Language. *Folia Linguistica* 32(1-2). doi:10.1515/flin.1998.32.1-2.23.

Benor, Sarah Bunin. 2010. Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity. *Journal of Sociolinguistics* 14(2). 159–183. doi:10.1111/j.1467-9841.2010.00440.x.

Berns, Jan. 2002. *Amsterdams*. (Taal in Stad En Land). Den Haag: SDU Uitgevers.

Bills, Garland D. 1977. Vernacular Chicano English: Dialect or Interference? *Journal of the Linguistic Association of the Southwest* 2. 30–36.

Boberg, Charles. 2004. Ethnic patterns in the phonetics of Montreal English. *Journal of Sociolinguistics* 8(4). 538–568. doi:10.1111/j.1467-9841.2004.00273.x.

Boersma, Paul. 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *IFA Proceedings* 17. 97–110.

Boersma, Paul & David Weenink. 2010. *Praat: doing phonetics by computer*. www.praat.org.

Booij, Geert. 1995. *The Phonology of Dutch*. Oxford: Oxford University Press. http://www.hum2.leidenuniv.nl/booijge/publicaties.htm (14 June, 2010).

Borg, Ingwer & Patrick J.F. Groenen. 2005. *Modern Multidimensional Scaling: Theory and Applications*. 2nd [rev.] ed. (Springer Series in Statistics). New York: Springer.

Brouwer, Dédé. 1989. Gender variation in Dutch: A sociolinguistic study of Amsterdam speech. Dordrecht: Foris Publications.

Brouwer, Dédé & Roeland van Hout. 1984. Amsterdam vernacular: the analysis of style shifting and covariation. *Linguistics in the Netherlands*. 33–43.

Brugman, Hennie & Albert Russel. 2004. Annotating Multimedia/Multi-modal resources with ELAN. In Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa & Raquel Silva (eds.), *Proceedings of LREC 2004, Fourth International Conference on Language Resources and Evaluation*, 2065–2068. Paris: European Language Resources Association.

Bybee, Joan L. 2001. *Phonology and language use*. (Cambridge Studies in Linguistics 94). Cambridge: Cambridge University Press.

Bybee, Joan L. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82(4). 711–733. doi:10.1353/lan.2006.0186.

Carlock, Elizabeth & Wolfgang Wölck. 1981. A method for isolating diagnostic linguistic variables: The Buffalo ehtnolects experiment. In David Sankoff & Henrietta Cedergren (eds.), *Variation Omnibus*., 17–24. Edmonton, Alberta: Linguistic Research, Inc. (Current Inquiry into Language, Linguistics and Human Communication 40.).

CBS. 2006a. *Begrippen*. Centraal Bureau voor de Statistiek (CBS; "Central Statistical Office"). http://www.cbs.nl/nl-NL/menu/methoden/begrippen/default.htm? Concept ID =315.

CBS. 2006b. *Definitions*. Centraal Bureau voor de Statistiek (CBS; "Central Statistical Office"). http://www.cbs.nl/en-GB/menu/methoden/begrippen/default.htm?Languageswitch=on& ConceptID=315.

CBS. 2014. *Statline*. Den Haag/Heerlen: Centraal Bureau voor de Statistiek (CBS; "Central Statistical Office"). http://statline.cbs.nl/ (18 December, 2014).

Charry, Eddy. 1983. Een sociolinguistische verkenning van het Surinaams-Nederlands [A sociolinguistic exploration of Suriname Dutch]. In Eddy Charry, Geert Koefoed & Pieter Muysken (eds.), *De taalen van Suriname: achtergronden en ontwikkelingen*, 138–161. Muiderberg.

Cheshire, Jenny, Sue Fox, Paul Kerswill & Eivind Torgersen. 2008. Ethnicity, friendship network and social practices as the motor of dialect change: Linguistic innovation in London. In Ulrich Ammon & Klaus Mattheier (eds.), *Sociolinguistica*, vol. 22, 1–23. http://www.degruyter.com/view/j/soci.2008.22.issue-1/9783484605299.1/9783484605299.1.xml?format=INT.

Cheshire, Jenny, Paul Kerswill, Sue Fox & Eivind Torgersen. 2011. Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics* 15(2). 151–196. doi:10.1111/j.1467-9841. 2011.00478.x.

Clyne, Michael. 2000. Lingua Franca and ethnolects in Europe and beyond. *Sociolinguistica* 14. 83–89.

Clyne, Michael. 2003. *Dynamics of language contact: English and immigrant languages*. (Cambridge Approaches to Language Contact). Cambridge ; New York: Cambridge University Press.

Cohen, Antonie, Carl L. Ebeling, Klaas Fokkema & André van Holk. 1971. Fonologie van het Nederlands en het Fries. Inleiding tot de moderne klankleer [Phonology of Dutch and Frisian. Introduction to the modern phonology]. 's-Gravenhage: Nijhoff.

Commissie Spreiding en Bereikbaarheid Nijmegen. 2003. Beleidsnotitie 28 april 2003. De bereikbaarheid en spreiding van de nijmeegse basisscholen en de samenstelling van de schoolpopulatie. Nijmegen: Gemeente Nijmegen. http://www2.nijmegen.nl/mmbase/attachments/135114/beleidsnotitie_28_april_20031.pdf.

Comrie, Bernard. 1997. Turkish phonology. In Alan S Kaye (ed.), *Phonologies of Asia and Africa. Volume 2*, 883–898. WinonaLake, IN: Eisenbrauns.

Cornips, Leonie. 2002. Etnisch Nederlands in Lombok [Ethnic Dutch in Lombok]. In Hans Bennis, Guus Extra, Pieter Muysken & Jacomine Nortier (eds.), *Een buurt in beweging. Talen en culturen in het Utrechtse Lombok en Transvaal*, 285 – 302. Amsterdam: Stichting Beheer IISG.

Cornips, Leonie. 2008. Loosing grammatical gender in Dutch: The result of bilingual acquisition and/or an act of identity? *International Journal of Bilingualism* 12(1-2). 105–124. doi:10.1177/13670069080120010701.

Cutler, Cecilia. 2008. Brooklyn style: Hip-hop markers and racial affiliation among European immigrants in New York City. *International Journal of Bilingualism* 12(1-2). 7–24. doi:10.1177/13670069080120010201.

Daan, Johanna Catharina. 1948. *Hij zeit wat: grepen uit de Amsterdamse volkstaal*. Amsterdam: Jacob van Campen.

Danesi, Marcel. 1985. A glossary of lectal terms for the description of language variation. *Language Problems and Language Planning* 89(2). 115–124.

den Besten, Hans & Frans Hinskens. 2005. Diversificatie van het Nederlands door taalcontact [Diversification of Dutch through language contact]. *Nederlandse Taalkunde* 10(4). 283–309.

Deppermann, Arnulf. 2007. Playing with the voice of the other: Stylized Kanakspak in conversations among German adolescents. In Peter Auer (ed.), *Style and Social Identities. An Alternative Approach to Linguistic Heterogeneity*, 325–360. Berlin/New York: De Gruyter.

de Vries, Jan W. 2005. Indisch-Nederlands. In Nicoline van der Sijs (ed.), *Wereldnederlands: Oude en Jonge Variëteiten van het Nederlands*, 59–78. Den Haag: Sdu Uitgevers. http://dbnl.nl/arch/sijs002were01_01/pag/sijs002were01_01.pdf.

Dirim, İnci & Peter Auer. 2004. Türkisch sprechen nicht nur die Türken. Über die Unschärfebeziehung zwischen Sprache und Ethnie in Deutschland. Berlin / New York: de Gruyter.

Dorleijn, Margreet & Jacomine Nortier. 2006. "Het is een grappige accent, weet je." Het opvallende prestige van de Marokkaanse tongval ["It is a funny accent, you know." The striking prestige of the Moroccan accent]. *Onze Taal op internet*. http://onzetaal.nl/2006/23/marokkaans.php.

Eckert, Penelope. 2008a. Where do ethnolects stop? *International Journal of Bilingualism* 12(1-2). 25–42. doi:10.1177/13670069080120010301.

Eckert, Penelope. 2008b. Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 453–476. doi:10.1111/j.1467-9841.2008.00374.x.

Eckert, Penelope. 2012. Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. *Annual Review of Anthropology* 41(1). 87–100. doi:10.1146/annurev-anthro-092611-145828.

El-Aissati, Abderrahman, Louis Boumans, Leonie Cornips, Margreet Dorleijn & Jacomine Nortier. 2005. Turks- en Marokkaans-Nederlands [Turkish and Moroccan Dutch]. In Nicoline van der Sijs (ed.), *Wereldnederlands: Oude en Jonge Variëteiten van het Nederlands*, 149–183. Den Haag: Sdu Uitgevers. http://dbnl.nl/arch/sijs002were01_01/pag/sijs002were01_01.pdf.

El-Aissati, Abderrahman & Yahya E-rramdani. 2001. Berbers. In Guus Extra & Jan Jaap de Ruiter (eds.), *Babylon aan de Noordzee. Nieuwe talen in Nederland*, 60–77. Amsterdam: Bulaaq.

Fagyal, Zsuzanna. 2005. Prosodic consequences of being a Beur: French in contact with immigrant languages in Paris. *Selected Papers from NWAV 32, Philadelphia, Working Papers in Linguistics* 10(2). 91–104.

Fagyal, Zsuzanna & Christopher M. Stewart. 2011. Prosodic style-shifting in preadolescent peer-group interactions in a working-class suburb of Paris. In Friederike Kern & Margret Selting (eds.), *Ethnic Styles of Speaking in European Metropolitan Cities*, 75–99. (Studies in Language Variation 8). Amsterdam / Philadelphia: Benjamins.

Fox, Sue, Arfaan Khan & Eivind Torgersen. 2011. The emergence and diffusion of Multicultural English. In Friederike Kern & Margret Selting (eds.), *Ethnic Styles of Speaking in European Metropolitan Cities*, 19–44. (Studies in Language Variation 8). Amsterdam / Philadelphia: Benjamins.

Fraurud, Kari & Sally Boyd. 2006. The native / non-native speaker distinction and the diversity of linguistics of young people in Swedish multilingual urban contexts. In Frans Hinskens (ed.), *Studies in Language Variation*, vol. 1, 53–69. Amsterdam: John Benjamins Publishing Company.

Fraurud, Kari & Sally Boyd. 2011. The native–non-native speaker distinction and the diversity of linguistic profiles of young people in multilingual urban contexts in Sweden. In Roger Källström & Inger Lindberg (eds.), *Young urban Swedish. Variation and change in multilingual settings*, 67–87. Gothenburg: University of Gothenburg. http://gupea.ub.gu.se/bitstream/2077/26570/1/gupea_2077_26570_1.pdf.

Freywald, Ulrike, Katharina Mayr, Tiner Özçelik & Heike Wiese. 2011. Kiezdeutsch as a multiethnolect. In Friederike Kern & Margret Selting (eds.), *Ethnic Styles of Speaking in European Metropolitan Cities*, 45–73. (Studies in Language Variation 8). Amsterdam / Philadelphia: Benjamins.

Frisch, Stefan A. 2011. Frequency effects. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*, vol. IV: Phonological interfaces, 2137–2163. Malden, MA: Wiley-Blackwell.

Geval, Tim. 2011. Etnische variëteiten in het Nederlands. “Leuk”, ‘opvallend, “apart” en andere indrukken uit het Nederlandse taalgebied. Twee onderzoeken naar de indrukken die niet-linguïstisch geschoolden hebben van Turks- en Marokkaans Nederlands. Utrecht: Universiteit Utrecht M.A. thesis.

Göksel, Aslı & Celia Kerslake. 2005. *Turkish: a comprehensive grammar*. London: Routledge.

Goossens, Jan, Johan Taeldeman & G. Verleyen. 2000. *Fonologische Atlas van de Nederlandse Dialecten*. Vol. III De Westgermaanse lange vocalen en diftongen. Gent: Koninklijke Academie voor Nederlandse Taal- en Letterkunde.

Grondelaers, Stefan, Paul van Gent & Roeland van Hout. 2015. Is Moroccan-flavoured Standard Dutch standard or not? On the use of perceptual criteria to determine the limits of standard languages. In Alexei Prikhodkine & Dennis R. Preston (eds.), *Responses to Language Varieties: Variability, processes and outcomes*, 191–220. (IMPACT: Studies in Language and Society 39). Amsterdam: John Benjamins Publishing Company.

Gussenhoven, Carlos. 1999. Illustrations of the IPA: Dutch. *Handbook of the International Phonetic Association*, 74–77. Cambridge: Cambridge University Press.

Guy, Gregory R. 2013. The cognitive coherence of sociolects: How do speakers handle multiple sociolinguistic variables? *Journal of Pragmatics* 52. 63–71. doi:10.1016/j.pragma.2012.12.019.

Harrell, Richard Slade. 1962. *A short reference grammar of Moroccan Arabic*. Washington, DC: Georgetown University Press.

Harrell, Richard Slade. 1965. *A basic course in Moroccan Arabic*. Washington, DC: Georgetown University Press.

Harris, Roxy. 2006. *New Ethnicities and Language Use*. (Ed.) Sue Wright & Helen Kelly-Holmes. Basingstoke: Palgrave Macmillan. http://www.palgraveconnect.com/doifinder/10.1057/9780230626461.

Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. (Routledge Curzon Arabic Linguistics Series). London: Routledge & Curzon.

Hewitt, Roger. 1986. *White talk black talk: inter-racial friendship and communication amongst adolescents*. (Comparative Ethnic and Race Relations). Cambridge [Cambridgeshire] ; New York: Cambridge University Press.

Hinskens, Frans. 1992. Dialect levelling in Limburg. Structural and sociolinguistic aspects. Nijmegen, The Netherlands: University of Nijmegen Ph.D. dissertation.

Hinskens, Frans. 1995. Dialectnivellering in Limburg: het suffix -lijk in Rimburg. *Gramma/TTT* 4(1). 49–70.

Hinskens, Frans. 1996. Nacht in Kerkrade. Dorsale fricatief deletie en woordfinale t-deletie in Ripuarische dialecten van het Nederlands. In Roeland van Hout & Joep Kruijsen (eds.), *Taalvariaties. Toonzettingen en modulaties op een thema. Feestbundel ter gelegenheid van de zestigste verjaardag van Toon Hagen*, 127–142. Dordrecht: Foris.

Hinskens, Frans. 2004. Nieuwe regenboogkleuren: jonge typen niet-standaardtaal en hun taalkundig belang. Amsterdam: VU Boekhandel/Uitgeverij.

Hinskens, Frans. 2011. Emerging Moroccan and Turkish varieties of Dutch: ethnolects or ethnic styles? In Friederike Kern & Margret Selting (eds.), *Ethnic Styles of Speaking in European Metropolitan Areas*, 101–129. (Studies in Language Variation 8). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Hinskens, Frans. 2013. Four decades of study of synchronic variation in varieties of Dutch. A sketch. In Joana Duarte & Ingrid Gogolin (eds.), *Linguistic superdiversity in urban areas: research approaches*, 227–252. (Hamburg Studies on Linguistic Diversity 2). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Hinskens, Frans. 2014. Despite or because of intensive contact?: Internal, external and extralinguistic aspects of divergence in modern dialects and ethnolects of Dutch. In Kurt Braunmüller, Steffen Höder & Karoline Kühl (eds.), *Stability and Divergence in Language Contact. Factors and Mechanisms*, 109–140. (Studies in Language Variation 16). Amsterdam: John Benjamins Publishing Company.

Hinskens, Frans, Peter Auer & Paul Kerswill. 2005. The study of dialect convergence and divergence: conceptual and methodological considerations. In Peter Auer, Frans Hinskens & Paul Kerswill (eds.), *Dialect Change. Convergence and Divergence in European Languages*, 1–48. Cambridge: Cambridge University Press.

Hinskens, Frans & Pieter Muysken. 2007. The talk of the town: languages in Amsterdam 1507-2007. (Ed.) Rita Franceschini. *Zeitschrift für Literaturwissenschaft und Linguistik (LiLi)* 37(148) [Special Issue: *Im Dickicht der Städte I: Sprache und Semiotik]*. 7–23.

Hinskens, Frans, Roeland van Hout, Pieter Muysken & Ariën van Wijngaarden. in preparation. Variation in the expression of grammatical gender in modern ethnolectal Dutch.

Hoffman, Michol F. & James A. Walker. 2010. Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change* 22(01). 37–67. doi:10.1017/S0954394509990238.

Horvath, Barbara & David Sankoff. 1987. Delimiting the Sydney speech community. *Language in Society* 16(02). 179–204. doi:10.1017/S0047404500012252.

Jacobi, Irene. 2009. On Variation and Change in Diphthongs and Long Vowels of Spoken Dutch. Universiteit van Amsterdam Ph.D. dissertation. http://dare.uva.nl/record/293297.

Jamin, Mikaël. 2004. “Beurs” and accent des cités: a case study of linguistic diffusion in la Courneuve. *Contemporary French and Francophone Studies* 8(2). 169–176. doi:10.1080/1026021042000199355.

Jamin, Mikaël. 2005. Sociolinguistic variation in the Parissuburbs. University of Kent at Canterbury Ph.D. dissertation.

Jamin, Mikaël, Cyril Trimaille & Médéric Gasquet-Cyrus. 2006. De la convergence dans la divergence: le cas des quartiers pluri-ethniques en France. *Journal of French Language Studies* 16(03). 335–356. doi:10.1017/S0959269506002559.

Jaspers, Jürgen. 2006. Stylizing Standard Dutch by Moroccan boys in Antwerp. *Linguistics and Education* 17(2). 131–156. doi:10.1016/j.linged.2006.09.001.

Jaspers, Jürgen. 2008. Problematizing ethnolects: Naming linguistic practices in an Antwerp secondary school. *International Journal of Bilingualism* 12(1-2). 85–103. doi:10.1177/13670069080120010601.

Jaspers, Jürgen. 2011. Strange bedfellows: Appropriations of a tainted urban dialect. *Journal of Sociolinguistics* 15(4). 493–524. doi:10.1111/j.1467-9841.2011. 00502.x.

Jaspers, Jürgen & Henk Aertsen. 2004. Marokkaanse jongens en het Antwerps dialect [Moroccan boys and the Antwerp dialect]. In Arjan van Leuvensteijn & Roeland van Hout (eds.), *Taalvariatie en groepsidentiteit*, 135–165. (Taal En Tonggeval Themanummer 17). Amsterdam: Vrije Universiteit.

Jørgensen, Jens Normann. 2005. Plurilingual conversations among bilingual adolescents. *Journal of Pragmatics* 37(3). 391–402. doi:10.1016/j.pragma.2004. 10.009.

Kallmeyer, Werner & Inken Keim. 2003. Linguistic variation and the construction of social identity in a German-Turkish setting: A case study of an immigrant youth group in Mannheim, Germany. In Jannis K. Androutsopoulos & Alexandra Georgakopoulou (eds.), *Discourse Constructions of Youth Identities*, 29–46. (Pragmatics & Beyond New Series 110). Amsterdam: John Benjamins Publishing Company.

Källström, Roger & Inger Lindberg (eds.). 2011. *Young urban Swedish. Variation and change in multilingual settings*. Gothenburg: University of Gothenburg. http://gupea.ub.gu.se/bitstream/2077/26570/1/gupea_2077_26570_1.pdf (20 May, 2015).

Kehrein, Roland. 2012. *Regionalsprachliche Spektren im Raum. Zur linguistischen Struktur der Vertikale*. (Zeitschrift Für Dialektologie Und Linguistik – Beihefte Bd. 152). Stuttgart: Franz Steiner Verlag.

Keim, Inken. 2002. Bedeutungskonstitution und Sprachvariation. Funktionen des "Gastarbeiterdeutsch" in Gesprächen jungendlicher Migrantinnen. In Arnulf Deppermann & Thomas Spranz-Fogasy (eds.), *Be-deuten. Wie Bedeutung im Gespräch entsteht*, 134–157. Tübingen: Stauffenburg.

Keim, Inken. 2007. Formen und Funktionen von Ethnolekten in multilingualen Lebenswelten - am Beispiel von Mannheim. (Ed.) Rita Franceschini. *Zeitschrift für Literaturwissenschaft und Linguistik (LiLi)* 37(148) [Special Issue: *Im Dickicht der Städte I: Sprache und Semiotik]*. 7–23.

Keim, Inken & Ralf Knöbl. 2007. Sprachliche Varianz und sprachliche Virtuosität türkisch-stämmiger Ghetto-Jugendlicher in Mannheim [Linguistic variance and linguistic virtuosity of ghetto-adolescents of Turkish origin in Mannheim]. In Christian Fandrych & Reinier Salverda (eds.), *Standard, Variation und Sprachwandel in germanischen Sprachen / Standard, Variation and Language Change in Germanic Languages*, 157–199. Tübingen: Narr.

Keim, Inken & Ralf Knöbl. 2011. Linguistic variation and linguistic virtuosity of young "Ghetto"-migrants in Mannheim, Germany. In Friederike Kern & Margret Selting (eds.), *Ethnic Styles of Speaking in European Metropolitan Cities*, 239–263. (Studies in Language Variation 8). Amsterdam / Philadelphia: Benjamins.

Kern, Friederike. 2011a. Rhythm in Turkish German talk-in-interaction. In Friederike Kern & Margret Selting (eds.), *Ethnic Styles of Speaking in European Metropolitan Cities*, 161–190. (Studies in Language Variation 8). Amsterdam / Philadelphia: Benjamins.

Kern, Friederike. 2011b. Introduction. In Friederike Kern & Margret Selting (eds.), *Ethnic Styles of Speaking in European Metropolitan Cities*, 1–18. (Studies in Language Variation 8). Amsterdam / Philadelphia: Benjamins.

Kern, Friederike. 2015. Turkish German. *Language and Linguistics Compass* 9(5). 219–233. doi:10.1111/lnc3.12136.

Kern, Friederike & Margret Selting (eds.). 2011. *Ethnic styles of speaking in European Metropolitan cities*. (Studies in Language Variation 8). Amsterdam/Philadelphia: Benjamins.

Kerswill, Paul, Eivind Torgersen & Sue Fox. 2008. Reversing "drift": Innovation and diffusion in the London diphthong system. *Language Variation and Change* 20(03). 451. doi:10.1017/S0954394508000148.

Khan, Arfaan. 2006. A sociolinguistic study of Birmingham English: Language variation and change in a multi-ethnic British community. Lancaster: Lancaster University Ph.D. dissertation.

Kiliç, Mehmet Akif & Fatih Öğüt. 2004. A high unrounded vowel in Turkish: is it a central or back vowel? *Speech Communication* 43(1-2). 143–154. doi:10.1016/j.specom.2004.03.001.

Kiparsky, Paul. 1973. "Elsewhere" in phonology. In Stephen R. Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 93–106. New York: Holt, Reinhart, and Winston.

Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.

Kossmann, Maarten G & Harry J Stroomer. 1997. Berber phonology. In Alan S Kaye (ed.), *Phonologies of Asia and Africa. Volume 1*, 461–475. WinonaLake, IN: Georgetown University Press.

Kotsinas, Ulla-Britt. 1988. Immigrant children's Swedish — A new variety? *Journal of Multilingual and Multicultural Development* 9(1-2). 129–140. doi:10.1080/01434632.1988.9994324.

Kotsinas, Ulla-Britt. 1998. Language contact in Rinkeby, an immigrant suburb. In Jannis Androutsopoulos & Arno Scholz (eds.), *Jugendsprache – langue des jeunes – youth language*, 125–148. Frankfurt etc.: Lang.

Koutsoudas, Andreas, Gerald Sanders & Craig Noll. 1974. The Application of Phonological Rules. *Language* 50(1). 1–28. doi:10.2307/412007.

Kuijpers, Erika. 2005. *Migrantenstad: immigratie en sociale verhoudingen in 17e-eeuws Amsterdam*. (Amsterdamse historische reeks d. 32). Hilversum: Verloren.

Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

Labov, William. 1971. The study of language in its social context. In Joshua A Fishman (ed.), *Advances in the sociology of language*, vol. Vol. I, Basic concepts, theories and problems: alternative approaches., 152–216. The Hague; Paris: Mouton.

Labov, William. 1972a. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William. 1972b. *Language in the inner city: studies in the Black English vernacular*. Philadelphia: University of Pennsylvania Press.

Labov, William. 2001. Principles of Linguistic Change, Volume II, Social Factors. Oxford: Blackwell.

Labov, William. 2014. The role of African Americans in Philadelphia sound change. *Language Variation and Change* 26(01). 1–19. doi:10.1017/S0954394513000240.

Ladefoged, Peter. 2005. Turkish (Istanbul dialect). *UCLA Phonetics Lab Data*. http://www.phonetics.ucla.edu/appendix/languages/turkish/turkish.html (18 August, 2011).

Lafkioui, Mena. 2007. Atlas linguistique des variétés du Rif [Linguistic atlas of variants of the Rif] (Berber Studies 16). Köln: Rüdiger Köppe Verlag.

Lameli, Alfred. 2004. *Standard und Substandard. Regionalismen im diachronen Längsschnitt*. (Zeitschrift Für Dialektologie Und Linguistik – Beihefte bd. 128). Stuttgart: Franz Steiner Verlag.

Lenz, Alexandra N. 2003. *Struktur und Dynamik des Substandards. Eine Studie zum Westmitteldeutschen (Wittlich/Eifel)*. (Zeitschrift Für Dialektologie Und Linguistik – Beihefte bd. 125). Stuttgart: Franz Steiner Verlag.

Lewis, Geoffrey. 2000[1967]. *Turkish grammar*. 2nd ed. Oxford: Oxford University Press.

Maddieson, Ian & Kristin Precoda. 2011. *The UCLA Phonological Segment Inventory Database (UPSID)*. http://www.linguistics.ucla.edu/faciliti/sales/software.htm (1 July, 2011).

Ma, Roxana & Eleanor Herasimchuk. 1971. The linguistic dimensions of a bilingual neighborhood. *Bilingualism in the barrio*, 347–464. (Indiana University Publications. Language Science Monographs v. 7). Bloomington: Indiana University.

Ma, Roxana & Eleanor Herasimchuk. 1972. Speech styles in Puerto Rican bilingual speakers: a factor analysis of co-variation of phonological variables. In Joshua A Fishman (ed.), *Advances in the sociology of languages*, vol. II, 268–295. The Hague: Mouton.

Marzo, Stefania & Evy Ceuleers. 2011. The use of Citétaal among adolescents in Limburg: the role of space appropriation in language variation and change. *Journal of Multilingual and Multicultural Development* 32(5). 451–464. doi:10.1080/01434632.2011.597560.

Mazzella, Léon. 2005. Le parler pied-noir: mots et expressions de là-bas. Paris: Rivages.

McClelland III, Clive W. 2008. *A phonology of Tarifit Berber*. Muenchen: Lincolm GmbH.

Mesthrie, Rajend. 2002. *Language in South Africa*. Cambridge, U.K.; New York: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511486692.

Mesthrie, Rajend. 2012. Ethnicity, substrate and place: The dynamics of Coloured and Indian English in five South African cities in relation to the variable (t). *Language Variation and Change* 24(03). 371–395. doi:10.1017/S0954394512000178.

Meunier, André. 2001. Remarques sur le français des pieds-noirs d'Oran [Remarks on the French language spoken by Oran Pieds-Noirs]. In K. Boucher (ed.), *Le Français et ses usages à l'écrit et à l'oral. Dans le sillage de S. Lafage*, 131–145. Paris: Presses de la Sorbonne nouvelle.

Mufwene, Salikoko S. 2001. *The ecology of language evolution*. Cambridge; New York: Cambridge University Press.

Muysken, Pieter. 2010. Ethnolects as a multidimensional phenomenon. In Muriel Norde, Bob de Jonge & Cornelius Hasselblatt (eds.), *Language Contact: New perspectives*, 7–26. Amsterdam: John Benjamins.

Muysken, Pieter. 2013. Ethnolects of Dutch. In Frans Hinskens & Johan Taeldeman (eds.), *Language and Space: Dutch*, 739–761. Berlin: Mouton de Gruyter. doi: 10.1515/ 9783110261332.739.

Nortier, Jacomine. 2002. "Tezz, wat is dat? Ik vind echt niet leuk ofzo!" Murks en straattaal. *Levende Talen Tijdschrift* 3(1). 30–36.

Nortier, Jacomine. 2008. Ethnolects? The emergence of new varieties among adolescents. *International Journal of Bilingualism* 12(1-2). 1–5. doi:10.1177/13670069080120010101.

Nortier, Jacomine & Margreet Dorleijn. 2008. A Moroccan accent in Dutch: A sociocultural style restricted to the Moroccan community? *International Journal of Bilingualism* 12(1-2). 125–142. doi:10.1177/13670069080120010801.

Nortier, Jacomine & Bente A. Svendsen (eds.). 2015. *Language, youth and identity in the 21st century: linguistic practices across urban spaces*. Cambridge: Cambridge University Press.

O+S Amsterdam. 2005. *Amsterdam in 15 stadsdelen en 97 buurtcombinaties, 1 januari 2005*. (Publicaties Gebiedsindeling Amsterdam). Amsterdam: Dienst/Bureau Onderzoek en Statistiek (O+S), gemeente Amsterdam. http://www.os.amsterdam.nl/popup/355 (31 December, 2014).

O+S Amsterdam. 2006. 5 Bevolking naar herkomstgroepering, 1 januari 2006. *Stadsdelen in cijfers 2006*, 22–24. Amsterdam: Dienst/Bureau Onderzoek en Statistiek (O+S), gemeente Amsterdam. http://www.os.amsterdam.nl/ (31 December, 2014).

O+S Amsterdam. 2007. 5 Bevolking naar herkomstgroepering, 1 januari 2007. *Stadsdelen in cijfers 2007*, 22–24. Amsterdam: Dienst/Bureau Onderzoek en Statistiek (O+S), gemeente Amsterdam. http://www.os.amsterdam.nl/ (31 December, 2014).

O+S Amsterdam. 2008. Tegengaan van segregatie in het basisonderwijs. Mogelijkheden en beperkingen van het instellen van voedingsgebieden. Amsterdam: Dienst/Bureau Onderzoek en Statistiek (O+S), gemeente Amsterdam.

O+S Amsterdam. 2009. 2.1.7 Bevolking naar herkomstgroepering, 1 januari 2005-2009. *Amsterdam in cijfers 2009*, 56. Amsterdam: Dienst/Bureau Onderzoek en Statistiek (O+S), gemeente Amsterdam. http://www.os.amsterdam.nl/feiten-en-cijfers/ (24 February, 2011).

O+S Amsterdam. 2010. *Stand van zaken segregatie basisonderwijs. Schooljaar 2009/'10*. (Ed.) Jolijn Broekhuizen & Esther Jakobs. Amsterdam: Dienst/Bureau Onderzoek en Statistiek (O+S), gemeente Amsterdam.

O+S Amsterdam. 2011. 2.1.5 Bevolking naar herkomstgroepering, 1 januari 2010. *Metropoolregio Amsterdam in cijfers 2010*, 35. Amsterdam: Dienst/Bureau Onderzoek en Statistiek (O+S), gemeente Amsterdam. http://www.os.amsterdam.nl/feiten-en-cijfers/ (28 December, 2014).

O+S Amsterdam. 2014. 2.1.7 Bevolking naar herkomstgroepering, 1 januari 2010-2014. *Amsterdam in cijfers 2014*, 62. Amsterdam: Dienst/Bureau Onderzoek en Statistiek (O+S), gemeente Amsterdam. http://www.os.amsterdam.nl/ (19 December, 2014).

Oostdijk, Nelleke, Wim Goedertier, Frank van Eynde, Louis Boves, Jean-Pierre Martens, Michael Moortgat & R. Harald Baayen. 2002. Experiences form the Spoken Dutch Corpus project. *Proceedings of the Third International Conference on Language Resources and Evaluation*, vol. 1, 340–347. Paris: ELRA.

O&S Nijmegen. 2003. *Leerlingenstromen en de samenstelling van schoolpopulaties*. 2nd ed. Nijmegen: Onderzoek en Statistiek (O&S), Gemeente Nijmegen. http://www2.nijmegen.nl/content/22524/eerder_verschenen_rapporten_ onderwijs.

O&S Nijmegen. 2004. *Onderwijsmonitor Nijmegen 2003*. Nijmegen: Onderzoek en Statistiek (O&S), Gemeente Nijmegen. http://www2.nijmegen.nl/content/22524/eerder_verschenen_rapporten_onderwijs.

O&S Nijmegen. 2014. *Stadsgetallen*. Nijmegen: Onderzoek en Statistiek (O&S), Gemeente Nijmegen. http://nijmegen.buurtmonitor.nl/ (18 December, 2014).

Pike, Kenneth Lee. 1945. *The intonation of American English*. Ann Arbor: University of Michigan Press.

Queen, Robin. 2012. Turkish-German bilinguals and their intonation: Triangulating evidence about contact-induced language change. *Language* 88(4). 791–816. doi:10.1353/lan.2012.0078.

Quist, Pia. 2000. New Copenhagen "Multi-ethnolect". Language use among adolescents in linguistic and culturally heterogeneous settings. *Danske Talesprog*, vol. 1, 142–212. København: C.A. Reitzels Forlag.

Quist, Pia. 2008. Sociolinguistic approaches to multiethnolect: Language variety and stylistic practice. *International Journal of Bilingualism* 12(1-2). 43–61. doi:10.1177/13670069080120010401.

Rampton, Ben. 1995. *Crossing: Language & Ethnicity Among Adolescents*. London & New York: Longman.

Rampton, Ben. 1998. Language crossing and the redefinition of reality: Expanding the agenda of research on code-switching. *Code-switching in conversation*, 290–317. London: Routledge.

Rampton, Ben. 2011. Style contrasts, migration and social class. *Journal of Pragmatics* 43(5). 1236–1250. doi:10.1016/j.pragma.2010.08.010.

Schatz, Henriëtte. 1986. Plat Amsterdams in its Social Context: A Sociolinguistic Study of the Dialect of Amsterdam. Amsterdam: P.J.Meertens-Instituut.

Schulte, A.G. 1982. Inleiding. *Het Rijk van Nijmegen. Westelijk gedeelte*, 4. Den Haag / Zeist: Staatsuitgeverij / Rijksdienst voor de Monumentenzorg. http://www.dbnl.org/tekst/schu211rijk02_01/schu211rijk02_01_0002.php.

Sebregts, Koen. 2015. The sociophonetics and phonology of Dutch "r." Utrecht: Utrecht University Ph.D. dissertation. http://www.lotpublications.nl/the-sociophonetics-and-phonology-of-dutch-r.

Selting, Margret. 2011. Prosody and unit construction in an ethnic style: the case of Turkish German and its use and function in conversation. In Friederike Kern & Margret Selting (eds.), *Ethnic Styles of Speaking in European Metropolitan Cities*, 130–160. (Studies in Language Variation 8). Amsterdam / Philadelphia: Benjamins.

Selting, Margret & Friederike Kern. 2009. On some syntactic and prosodic structures of Turkish German in talk-in-interaction. *Journal of Pragmatics* 41(12). 2496–2514. doi:10.1016/j.pragma.2009.05.018.

Şimşek, Yazgül. 2011. Constructions with the Turkish şey and its German equivalent dings in Turkish–German conversations. In Friederike Kern & Margret Selting (eds.), *Ethnic Styles of Speaking in European Metropolitan Cities*, 191–216. (Studies in Language Variation 8). Amsterdam / Philadelphia: Benjamins.

Smits, Caroline. 1996. Disintegration of inflection. The case of Iowa Dutch. Amsterdam: VU University Amsterdam.

Spiegel, Henrik L. 1962. Twe-spraack vande Nederduitsche letterkunst/ ófte Vant spellen ende eyghenscap des Nederduitschen taals. In W.J.H. Caron (ed.), *Twe-spraack. Ruygh-bewerp. Kort begrip. Rederijck-kunst*, 3–64. Groningen: J.B. Wolters. http://www.dbnl.org/tekst/spie001twes01_01/index.php.

Stroop, Jan. 1998. *Poldernederlands, waardoor het ABN verdwijnt*. Amsterdam: Bert Bakker.

Stroud, Christopher. 2004. Rinkeby Swedish and semilingualism in language ideological debates: A Bourdieuean perspective. *Journal of Sociolinguistics* 8(2). 196–214. doi:10.1111/j.1467-9841.2004.00258.x.

Stuart-Smith, Jane, Claire Timmins & Farhana Alam. 2011. Hybridity and ethnic accents: A sociophonetic analysis of "Glaswasian." In Frans Gregersen, Jeffrey K. Parrott & Pia Quist (eds.), *Language variation - European perspectives III*, vol. 7, 43–58. (Studies in Language Variation). Amsterdam: John Benjamins Publishing Company.

Svendsen, Bente Ailin & Unn Røyneland. 2008. Multiethnolectal facts and functions in Oslo, Norway. *International Journal of Bilingualism* 12(1-2). 63–83. doi:10.1177/13670069080120010501.

Swift, Lloyd B. 1963. *A reference grammar of Modern Turkish*. Bloomington: Indiana University Publications.

Tahitu, Bert. 1989. Melaju Sini, het Maleis van Molukse jongeren in Nederland [Melaju Sini, the Malay of Moluccan Youth in the Netherlands]. Leiden: Leiden University Ph.D. dissertation.

Thelwall, Robin & M. Akram Sa'Adeddin. 1990. Arabic. *Journal of the International Phonetic Association* 20(02). 37–39. doi:10.1017/S0025100300004266.

van de Berg, Marinus & Frank Pieke. 1991. De Chinezen. In Jan Jaap de Ruiter (ed.), *Talen in Nederland. Een beschrijving van de taalsituatie van negen etnische groepen*, 43–69. Groningen: Wolters-Noordhoff.

van der Harst, Sander. 2011. The Vowel Space Paradox: A Sociophonetic Study on Dutch. Nijmegen, The Netherlands: Radboud Universiteit Nijmegen Ph.D. dissertation. http://www.lotpublications.nl/the-vowel-space-paradox-the-vowel-space-paradox-a-sociophonetic-study-on-dutch.

van der Harst, Sander, Hans van de Velde & Roeland van Hout. 2014. Variation in Standard Dutch vowels: The impact of formant measurement methods on identifying the speaker's regional origin. *Language Variation and Change* 26(02). 247–272. doi:10.1017/S0954394514000040 (16 September, 2015).

van der Hulst, Harry & Jeroen van de Weijer. 1991. Topics in Turkish phonology. In Hendrik Boeschoten & Ludo Verhoeven (eds.), *Turkish Linguistics Today*, 11–59. Leiden: E.J. Brill.

van der Wal, Marijke J. 1992. Dialect and standard language in the past: "The rise of the Dutch standard language in the sixteenth and seventeenth centuries." In Jan Arnoldus van Leuvensteijn & J.B. Berns (eds.), *Dialect and standard language*, 119–129. Amsterdam / Oxford / New York / Tokyo: North-Holland.

van de Velde, Hans, Marinel Gerritsen & Roeland van Hout. 1996. The devoicing of fricatives in Standard Dutch: A real-time study based on radio recordings. *Language Variation and Change* 8. 149–175. doi:10.1017/S0954394500001125.

van Heuven, Vincent J., Loulou Edelman & Renée van Bezooijen. 2002. The pronunciation of /ɛi/ by male and female speakers of avant-garde Dutch. In H. Broekhuis & Paula Fikkert (eds.), *Linguistics in the Netherlands 2002*, 61–72. Amsterdam: John Benjamins.

van Heuven, Vincent J., Renée van Bezooijen & Loulou Edelman. 2005. Pronunciation of /ɛi/ in avant-garde Dutch: Across-sex acoustic study. In Markku Filppula, Juhani Klemola, Marjatta Palander & Esa Penttilä (eds.), *Dialects Across Borders*, 185–210. Amsterdam: John Benjamins.

van Heuven, Vincent J. & Els van Houten. 1985. De klinkers in het Nederlands van Turken [The vowels in the Dutch of Turks]. *Forum der Letteren* 26. 201–213.

van Hout, Roeland. 1989. De structuur van taalvariatie. Een sociolinguïstisch onderzoek naar het stadsdialect van Nijmegen [The structure of language variation. A sociolinguistic study of the city dialect of Nijmegen]. Dordrecht: Foris Publications.

van Hout, Roeland. 1999. De Nijmeegse stadstaal. Honderd jaar dialectontrouw aan de Waal. In Joep Kruijsen & Nicoline van der Sijs (eds.), *Honderd jaar stadstaal*, 75–85. Amsterdam/Antwerp: Contact.

van Hout, Roeland, Joep Kruijsen & Marinel Gerritsen. 2014. Exosmosis along the Romance-Germanic language border in Belgium. The diffusion of French borrowings in the Dutch dialects of Haspengouw. In Raquel Casesnoves-Ferrer, Montserrat Forcadell Guinjoan & Núria Gavaldà Ferré (eds.), *Ens queda la paraula: estudis de lingüística aplicada en honor a M. Teresa Turell*, 197–223. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.

van Hout, Roeland, Jan Roelofs & Leo van Stijn. 2006. *Kie(k) dääör, wà rääör: woordenboek van het Nimweegs dialect*. Nijmegen: Selexyz Dekker v.d. Vegt.

van Krieken, Esther. 2004. Van immigratie tot etnolect. Een onderzoek naar het ontstaan en bestaan van het Marokkaans Nederlands in Nijmegen. Nijmegen, The Netherlands: Radboud University Nijmegen M.A. thesis.

van Krieken, Esther. 2005. Van immigratie tot etnolect [From immigration to ethnolect]. *Toegepaste Taalwetenschap in Artikelen* 73. 63–71.

van Lier, Eva. 2005. Straattaal. *Neerlandica Extra Muros*(43). 12–26.

van Meel, Linda, Frans Hinskens & Roeland van Hout. 2013. Ethnolectal variation in the realization of /z/ by Dutch youngsters. *Zeitschrift für Dialektologie und Linguistik* 80(3). 297–325.

van Meel, Linda, Frans Hinskens & Roeland van Hout. 2014. Variation in the realization of /ɛi/ by Dutch youngsters: from local urban dialects to emerging ethnolects? *Dialectologia et Geolinguistica* 22. 46–74. doi:10.1515/dialect-2014-0004.

VVV Arnhem Nijmegen. 2014. Rijk van Nijmegen. *VVV Arnhem Nijmegen*. http://www.vvvarnhemnijmegen.nl/rijk-van-nijmegen (29 December, 2014).

Weinreich, Uriel. 1953. *Languages in contact: findings and problems*. The Hague: Mouton.

Weinreich, Uriel, William Labov & Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Directions for historical linguistics*, 97–195. Austin: University of Texas Press. http://www.utexas.edu/cola/centers/lrc/books/hist05.html.

Wiese, Heike. 2009. Grammatical innovation in multiethnic urban Europe: New linguistic practices among adolescents. *Lingua* 119(5). 782–806. doi:10.1016/j.lingua.2008.11.002.

Wiese, Heike. 2013. From feature pool to pond: The ecology of new urban vernaculars. *Working Papers in Urban Language and Literacies* 104. https://www.kcl.ac.uk/innovation/groups/ldc/publications/workingpapers/WP104-Wiese-2013-The-ecology-of-new-urban-vernaculars--2-.pdf (2 September, 2013).

Wölck, Wolfgang. 2002. Ethnolects – between bilingualism and urban dialect. In Li Wei, Jean-Marc Dewaele & Alex Housen (eds.), *Opportunities and challenges of bilingualism*, 157–170. Berlijn: Mouton de Gruyter.

Appendix 1:
Selection of international ethnolectal studies

Author	Year	Ethnolect(s) / (involved) language(s)	Groups	City / Region
Labov (William)	1972	"BEV" (AAVE)	"Black"* youngsters (i.e. young Afro-Americans)	USA, NY, New York City
Ma (Roxana) & Herasimchuk (Eleanor)	1972	English and Spanish	45 bilingual Puerto Rican speakers	USA, NJ, Jersey City
Bills (Garland)	1977	Vernacular Chicano English (VCE)	Chicanos	USA, the Southwest
Carlock (Elizabeth) & Wölck (Wolfgang)	1981	Buffalo English and various ethnolects	Buffalo citizens, 3 ethnic groups: German, Italian, Polish	USA, NY, City of Buffalo
Hewitt (Roger)	1986	'London Jamaican' creole	"young black Londoners"* (i.e. young Afro-Caribbean Londoners); "young whites"	EU, UK, South London
Kotsinas (Ulla-Britt)	1988	Rinkeby Swedish	children with different home languages	EU, Sweden, Stockholm
Rampton (Ben)	1995	English, Panjabi, Creole, stylized Asian English	teenagers of Indian, Pakistani, African Caribbean and Anglo descent	EU, UK, South Midlands
Kotsinas (Ulla-Britt)	1998	Rinkeby Swedish	teenagers and children with immigrant backgrounds	EU, Sweden, Stockholm

* Terminology / description as given by author in the study concerned.

Author	Year	Features / Variables	Material	Dominant Perspective / Approach
Labov (William)	1972	mainly: copula deletion; negative concord; range of phonological variables	observations	Structural-linguistic
Ma (Roxana) & Herasimchuk (Eleanor)	1972	14 phonological features	sociolinguistic interview	Structural-linguistic
Bills (Garland)	1977	syntactic, phonological, phonetic	observations	Structural-linguistic
Carlock (Elizabeth) & Wölck (Wolfgang)	1981	"potentially diagnostic features": prosodic and phonological features, features in discourse and grammar	attitude study: informal interviewing; speech stimuli; subjective reaction test	Structural-linguistic
Hewitt (Roger)	1986	a.o. creole words, idioms, phonological + prosodic + intonational features	observations; interviews	Ethnographic
Kotsinas (Ulla-Britt)	1988	pronunciation; vocabulary	interviews? observations?	Structural-linguistic
Rampton (Ben)	1995	crossing of words, expressions	fieldwork: a.o. radio-microphone recording, interviews, participant observations, retrospective discussion with participants on extracts of recordings	Ethnographic
Kotsinas (Ulla-Britt)	1998	pronunciation; vocabulary; morphology	teacher survey: judgements; recordings	Structural-linguistic

Author	Year	Ethnolect(s) / (involved) language(s)	Groups	City / Region
Rampton (Ben)	1998	English, Panjabi, Creole, stylized Asian English	teenagers of Indian, Pakistani, African Caribbean and Anglo descent	EU, UK, South Midlands
Keim (Inken)	2002	'Gastarbeiter-deutsch': German variety of 1st generation working migrants	“powergirls”: young immigrant women (2nd generation), almost all of Turkish origin	EU, Germany, Mannheim
Clyne (Michael)	2003	ethnolects of Australian English (e.g. Greek-, Yiddish-, German-based)	4 families of German, Greek and Hungarian descent; Italian-Australians; ultra-orthodox Jewish community	Australia, Melbourneand state of Victoria
Kallmeyer (Werner) & Keim (Inken)	2003	repertoire: some kind of dialectal variety of their country of origin; several German and language mixing varieties	15 teenage girls, most from Turkish origin, 2 of Italian, 1 African, 1 Bosnian, 1 Thai	EU, Germany, Mannheim

Author	Year	Features / Variables	Material	Dominant Perspective / Approach
Rampton (Ben)	1998	crossing of words, expressions	radio-microphone recording, participant observations, interviewing, retrospective participant commentary on extracts of recorded interaction	Ethnographic
Keim (Inken)	2002	use and function of 'Gastarbeiter-deutsch' (meta-pragmatic)	conversations	Ethnographic "one aspect of the communicative repertoire of a group of young migrant women [...] use this speech variety in specific sociosymbolic functions"
Clyne (Michael)	2003	o.a. phonetic, phonological, prosodic features; syntactic marking, lexical items	brief case studies	Structural-linguistic
Kallmeyer (Werner) & Keim (Inken)	2003	“variation practices, where elements of one language are locally combined with elements of the other”: o.a. grammatical connections: e.g. interjections; morphosyntactically integrated constructions; pragmatics: discourse functions	ethnographic interviews and participant observation	Ethnographic “part of an ethnographically based sociolinguistic study of various immigrant youth groups and their social style of communication”

Author	Year	Ethnolect(s) / (involved) language(s)	Groups	City / Region
Boberg (Charles)	2004	Montreal English, Canadian English	Montrealers of Irish, Italian, and Jewish ethnic origin	Canada, Montreal
Dirim (İnci) & Auer (Peter)	2004	Turkish German ethnolect	youth of Turkish and non-Turkish descent	EU, Germany, Hamburg
Jamin (Mikaël)	2004	*'accent des cités'*	2 age groups of 'Metropolitan', 'North African' and of 'other' origin	EU, France, Paris, neighborhood La Courneuve
Fagyal (Zsuzanna)	2005	*Berbère* or a vernacular variety of Arabic; French	male speakers of 12+14 years of age who were of North African descent (i.e. *'Beurs'* / 'French-Arabic') or Caucasian descent (i.e. 'French-French')	EU, France, Paris, neighborhood La Courneuve
Jørgensen (Jens)	2005	Danish; Turkish; range of varieties from English, French, German, to stylistic, stereotypical immigrant Swedish	Turkish-Danish grade school students	EU, Denmark
Harris (Roxy)	2006	English, Panjabi, Gujarati, Hindi, and Urdu	Blackhill Youth, i.e. adolescents of mainly South Asian descent, *Brasians*	EU, UK, West London

Author	Year	Features / Variables	Material	Dominant Perspective / Approach
Boberg (Charles)	2004	vowels	tape-recorded word list	Structural-linguistic
Dirim (İnci) & Auer (Peter)	2004	phonological features, morphological and syntactical features	recordings	Ethnographic
Jamin (Mikaël)	2004	Affrication of t, d, k, and g	observations?	Structural-linguistic "a 'variationist' approach"
Fagyal (Zsuzanna)	2005	intonation	interviews; picture naming task	Structural-linguistic
Jørgensen (Jens)	2005	words, expressions	group conversations	Ethnographic
Harris (Roxy)	2006	phonologically, grammatically and lexically linguistic markers of identity and affiliation	'extensive survey of language use'; 30 written accounts; audio recordings; intensive, but open-ended interviews	Ethnographic

Author	Year	Ethnolect(s) / (involved) language(s)	Groups	City / Region
Jamin (Mikaël), Trimaille (Cyril) & Gasquet-Cyrus (Médéric)	2006	(non-)standard forms of French	French young people in multiethnic and multilingual neighborhoods	EU, France, Grenoble, Paris area, and Marseille
Jaspers (Jürgen)	2006	Stylized Standard Dutch	Belgian adolescents of Moroccan descent	EU, Belgium, Antwerp
Khan (Arfaan)	2006	traditional and ethnic variants of English	White English; Black Caribbean; Pakistani	EU, UK, Birmingham
Deppermann (Arnulf)	2007	Stylized 'Kanaksprak'	peer-group of German adolescents	EU, Germany
Cheshire (Jenny), Fox (Sue), Kerswill (Paul) & Torgersen (Eivind)	2008	Multicultural London English (MLE)	Anglo vs. non-Anglo aged 16-19	EU, UK, London

Author	Year	Features / Variables	Material	Dominant Perspective / Approach
Jamin (Mikaël), Trimaille (Cyril) & Gasquet-Cyrus (Médéric)	2006	dental and velar plosives	interview, lecture	Studie 1: Structural-linguistic; studies 2 + 3: Ethnographic
Jaspers (Jürgen)	2006	stylized instances: semantic, (meta-) pragmatic	participant observation, interviewing, individual/ classroom (audio) recording, feedback-interviews on extracts from recordings	Ethnographic
Khan (Arfaan)	2006	2 vowels	interviews	Structural-linguistic
Deppermann (Arnulf)	2007	phonetics / pronunciation, phonology / voice quality, grammar, syntax and turn-design, lexis, semantics, phraseology (i.e. abusive slogans/ritual sayings)	23 sequences of Stylized 'Kanaksprak' from audio-recordings	Ethnographic "conversation analytic view which holds that people's social identities and the features associated with them are neither invariably fixed nor relevant for just any interaction."
Cheshire (Jenny), Fox (Sue), Kerswill (Paul) & Torgersen (Eivind)	2008	phonological features (short/long monophthongs, diphthongs, consonants), quotatives	recordings of informal speech	Structural-linguistic

Author	Year	Ethnolect(s) / (involved) language(s)	Groups	City / Region
Cutler (Cecilia)	2008	Brooklyn style, hip-hop-styled speech	Eastern European immigrants	USA, NY, New York City, borough Brooklyn
Eckert (Penny)	2008	Chicano English; Anglo English	Latino; white Anglo	USA, CA, Northern California
Jaspers (Jürgen)	2008	Belgian-Moroccan adolescents' vernacular Dutch	Belgian adolescents of Moroccan descent	EU, Belgium, Antwerp
Kerswill (Paul), Torgersen (Eivind) & Fox (Sue)	2008	London English	Anglo vs non-Anglo	EU, UK, London
Svendsen (Bente Ailin) & Røyneland (Unn)	2008	Norwegian; several heritage languages; multi-ethnolectal speech	Norwegian L1 speakers regardless of ethnic descent; focus on 2 participants of North African descent	EU, Norway, Oslo

Author	Year	Features / Variables	Material	Dominant Perspective / Approach
Cutler (Cecilia)	2008	hip-hop markers: phonological, paralinguistic (e.g. falsetto voice, suck teeth), morpho-syntactic, discourse & lexical	informal interviews	Ethnographic
Eckert (Penny)	2008	vowels	observations; F1/F2 measurements	Largely Ethnographic
Jaspers (Jürgen)	2008	grammar, phonology and prosody	incl. participation observation, interviews, individual / group recordings	Ethnographic "ethnolects be viewed as representations of particular ways of speaking that do not necessarily correspond to systematic linguistic practices"
Kerswill (Paul), Torgersen (Eivind) & Fox (Sue)	2008	vowels	recordings	Structural-linguistic
Svendsen (Bente Ailin) & Røyneland (Unn)	2008	lexical, prosody, grammar	focus on: 1 interview and 1 peer conversation	Largely Ethnographic “a structural / dialectological (linguistic description) and sociopragmatic / functional approach”

Author	Year	Ethnolect(s) / (involved) language(s)	Groups	City / Region
Quist (Pia)	2008	Copenhagen multi-ethnolect; standard Danish, different 1st languages: a.o. Berber, Turkish, Kurdish, Arabic, Danish, Urdu and Serbian	2nd generation immigrant youth	EU, Denmark, Copenhagen
Selting (Margret) & Kern (Friederike)	2009	Turkish German, aka Türkendeutsch / Türkenslang / 'Kanak sprach'	Turkish German	EU, Germany, Berlin
Wiese (Heike)	2009	Kiezdeutsch	young people in multiethnic neighbourhoods of Berlin	EU, Germany, Berlin
Cheshire (Jenny), Kerswill (Paul), Fox (Sue), Torgersen (Eivind)	2011	Multicultural London English (MLE)	Anglo vs. non-Anglo	EU, UK, London

Author	Year	Features / Variables	Material	Dominant Perspective / Approach
Quist (Pia)	2008	morphology; syntax; lexicon; pronunciation	self-recordings, group recordings, retrospective interviews	Ethnographic "a *variety perspectiv*e and a *stylistic practice perspective*"
Selting (Margret) & Kern (Friederike)	2009	TCU's (turn-constructional units); syntactic and prosodic features	telephone and face-to-face conversations	Ethnographic "[…] some of these structures and show how they are used as a resource […] overview of the structures of phenomena typical of the Turkish German style of speaking"
Wiese (Heike)	2009	grammatical and pragmatic features: grammatical reductions; productive elaborations	spontaneous speech samples; informal, conversation-like group-recordings; self-recordings of a 17-year-old of Kurdish-Arabic descent	Structural-linguistic
Cheshire (Jenny), Kerswill (Paul), Fox (Sue), Torgersen (Eivind)	2011	vowels, an innovative discourse-pragmatic feature, a morphosyntactic feature, a morphophonological feature	recordings	Structural-linguistic "a variationist approach"

Author	Year	Ethnolect(s) / (involved) language(s)	Groups	City / Region
Fagyal (Zsuzanna) & Stewart (Christopher M.)	2011	peer group-based interactional styles (prosodic style-shifting)	adolescents in a multi-ethnic working-class suburb	EU, France, Paris, neighborhood La Courneuve
Fox (Sue), Khan (Arfaan), Torgersen (Eivind)	2011	Multicultural English with social and regional variation; traditional English varieties	Study 1 (Tower Hamlets study): White British, Bangladeshi, Mixed race White British / Black Caribbean; Study 2 (Linguistic Innovators study): Anglo (relatively local roots) vs. non-Anglo; Study 3: see Khan 2006	EU, UK, London (Studies 1 and 2) and Birmingham (= Khan 2006)
Freywald (Ulrike), Mayr (Katharina), Özçelik (Tiner), Wiese (Heike)	2011	multi-ethnolect Kiezdeutsch	adolescents from a multiethnic and a monoethnic neighbourhood of Berlin	EU, Germany, Berlin
Keim (Inken) & Knöbl (Ralf)	2011	the communicative German repertoire	one 17-year-old young “Ghetto”-migrant	EU, Germany, Mannheim

Author	Year	Features / Variables	Material	Dominant Perspective / Approach
Fagyal (Zsuzanna) & Stewart (Christopher M.)	2011	phrase-final intonation contours	recordings of face-to-face interactions; peer-group interactions in guided interviews	Ethnographic "[...] pragmatic meaning [...] function as a micro-level style feature indexing common ground and in-group affiliation"
Fox (Sue), Khan (Arfaan), Torgersen (Eivind)	2011	vowels	participant observations, interviews	Structural-linguistic Study 1: "ethnographic study [...] makes use of participant observation to examine friendship networks as well as sociolinguistic / ethnographic interviews for data elicitation" Study 3: “explores the extent of inter-ethnic linguistic variation”
Freywald (Ulrike), Mayr (Katharina), Özçelik (Tiner), Wiese (Heike)	2011	syntactic, morphological, lexical levels	a recognition study, i.e. acceptability test	Structural-linguistic “the perspective of a variety”
Keim (Inken) & Knöbl (Ralf)	2011	linguistic features that convey the impression of "foreignness”, i.e. phonetic and grammatical features, greeting rituals	2 interviews with 1 17-year-old and his friends	Ethnographic

Author	Year	Ethnolect(s) / (involved) language(s)	Groups	City / Region
Marzo (Stefania) & Ceuleers (Evy)	2011	Citétaal	adolescents with parents of (1) Italian, (2) Belgian, (3) mixed origin (incl. 1 parent of Italian origin), (4) both other origin than Italian	EU, Belgium, Eastern part of Flanders: Limburg (focus on Genk)
Rampton (Ben)	2011	two pairs of contrasting styles – posh and Cockney, and Creole and Asian English	study 1: adolescents at a multi-ethnic secondary school; study 2: multi-ethnic adolescent peer groups	EU, UK, Study 1: London; Study 2: neighborhood Ashmead in the south Midlands of England
Stuart-Smith (Jane), Timmins (Claire) & Alam (Farhana)	2011	Glasgow Asian / Glaswasian; English and Punjabi	Study 1: English dominant Punjabi Glasgow Asian speakers and non-Asians; Study 2: 18 year old 'Cultural' and 'Religious' English-dominant Punjabi / Urdu trilingual girls	EU, UK, Scotland, Glasgow
Mesthrie (Rajend)	2012	2 varieties of English: viz. those of its Coloured and Indian communities	persons of Colored, Indian and Black descent	South Africa, five South African cities

Author	Year	Features / Variables	Material	Dominant Perspective / Approach
Marzo (Stefania) & Ceuleers (Evy)	2011	use of *Citétaal*	a quantitative questionnaire study using self-reports on a 5-point Likert scale and qualitative analysis of focus group interviews	Ethnographic "a stylistic practice perspective"
Rampton (Ben)	2011	grammatical and phonological stances associated with standard and vernacular speech	field observations; interviews	Ethnographic
Stuart-Smith (Jane), Timmins (Claire) & Alam (Farhana)	2011	vowels of FACE and GOAT, consonant /l/	acoustic analysis on data of 2 studies: Study 1 consisted of experimental elicited data, Study 2 consisted of recordings	Structural-linguistic Data of 2 studies were used. Study 1 "was experimental", study 2 "used the communities of practice (CoP) framework"
Mesthrie (Rajend)	2012	the variable (t) as well as (d) and (th)	interviews, word list	Structural-linguistic

Author	Year	Ethnolect(s) / (involved) language(s)	Groups	City / Region
Queen (Robin)	2012	German	8 2nd- and 3rd-generation Turkish-German bilingual adults	EU, Germany, Berlin and Freiburg
Wiese (Heike)	2013	new urban vernacular / dialect: Kiezdeutsch (lit. “(neighbour-) hood German”	speakers across migrant and non-migrant backgrounds in multilingual urban neighborhoods	EU, Germany, Berlin

Author	Year	Features / Variables	Material	Dominant Perspective / Approach
Queen (Robin)	2012	intonation patterns (terminal rises)	narrations of picture book in German	Structural-linguistic "their intonational grammars have developed in a way consistent with contact-related language change and the development of a new variety of German"
Wiese (Heike)	2013	bare local NPs; new word order options; ‘gib(t)s’ as an existential marker	a.o. data from the KiezDeutsch Korpus (KiDKo)	Structural-linguistic “a ‘dialect’ point of view”

Appendix 2:
Interview questions

The interview questions are reorganized according to the meta-data categories in our database. The meta-data categories are indicated in bold.

Name: _____
Age: _____
Birth date: _____

Language skills
Which languages do you speak well? _____

Linguistic profile
Which language(s) do you speak with:
- your father? _____
- your mother? _____
- your younger brother(s)? _____
- your older brother(s)? _____
- your younger sister(s)? _____
- your older sister(s)? _____
- your friends? _____
- others? _____

Which language(s) does your father speak? _____
Which language(s) does your mother speak? _____

Family structure
With whom do you live at home?
- your father? yes / no
- your mother? yes / no
- your younger brother(s)? yes / no
- your older brother(s)? yes / no
- your younger sister(s) yes / no
- your older sister(s) yes / no
- Someone else? yes / no
  (for example grandpa and grandma)
  - if yes, who? _____

Age family

How old is:

- your father? _____
- your mother? _____

If you have brothers or sisters: how old are they?

- your younger brother(s) _____
- your older brother(s) _____
- your younger sister(s) _____
- your older sister(s) _____

Origin parents

Native country father: _____
Native country mother: _____
Place of birth father: _____
Place of birth mother: _____

Friend structure

Are your friends from the same neighborhood? _____
Which ethnicities do your friends have? _____
Who are your friends? _____

Living history

Where were you born? _____
In which neighborhood do you live? _____
Since when? _____
Where else have you lived? _____

Other information

What are your hobbies? _____
_____
_____

Appendix 3: Maps

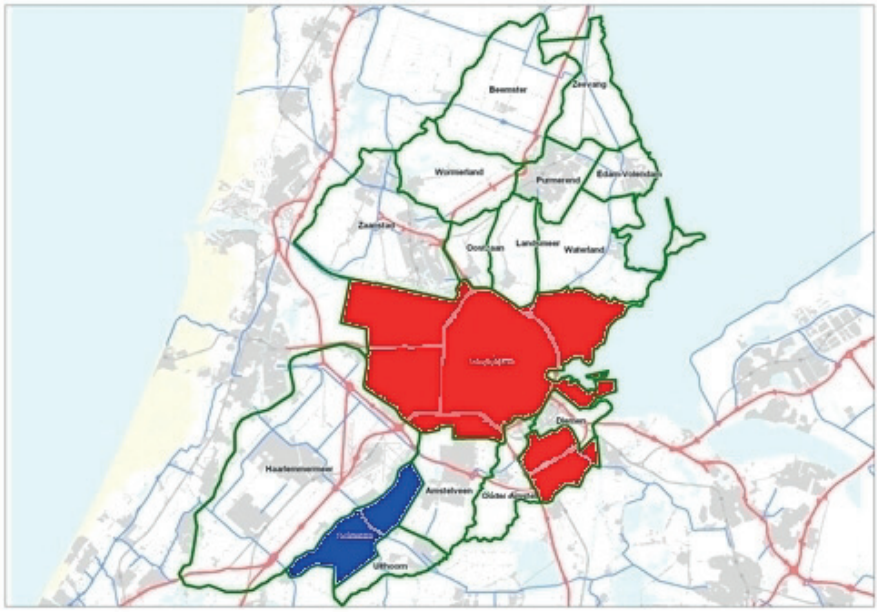
Map 1 *Stadsregio* Amsterdam 'City Region of Amsterdam, i.e. the metropolitan region of Amsterdam' with its 16 neighboring municipalities

Map 2 Amsterdam with its 15 *stadsdelen* 'boroughs' and 97 *buurtcombinaties* 'neighborhood combinations' before May 1st 2009

Map 3 Nijmegen with its 9 *stadsdelen* 'boroughs' and 44 *wijken* 'neighborhoods'

Map 4 *Rijk van Nijmegen,* lit. 'Land of Nijmegen' with the four municipalities and villages relevant to our study

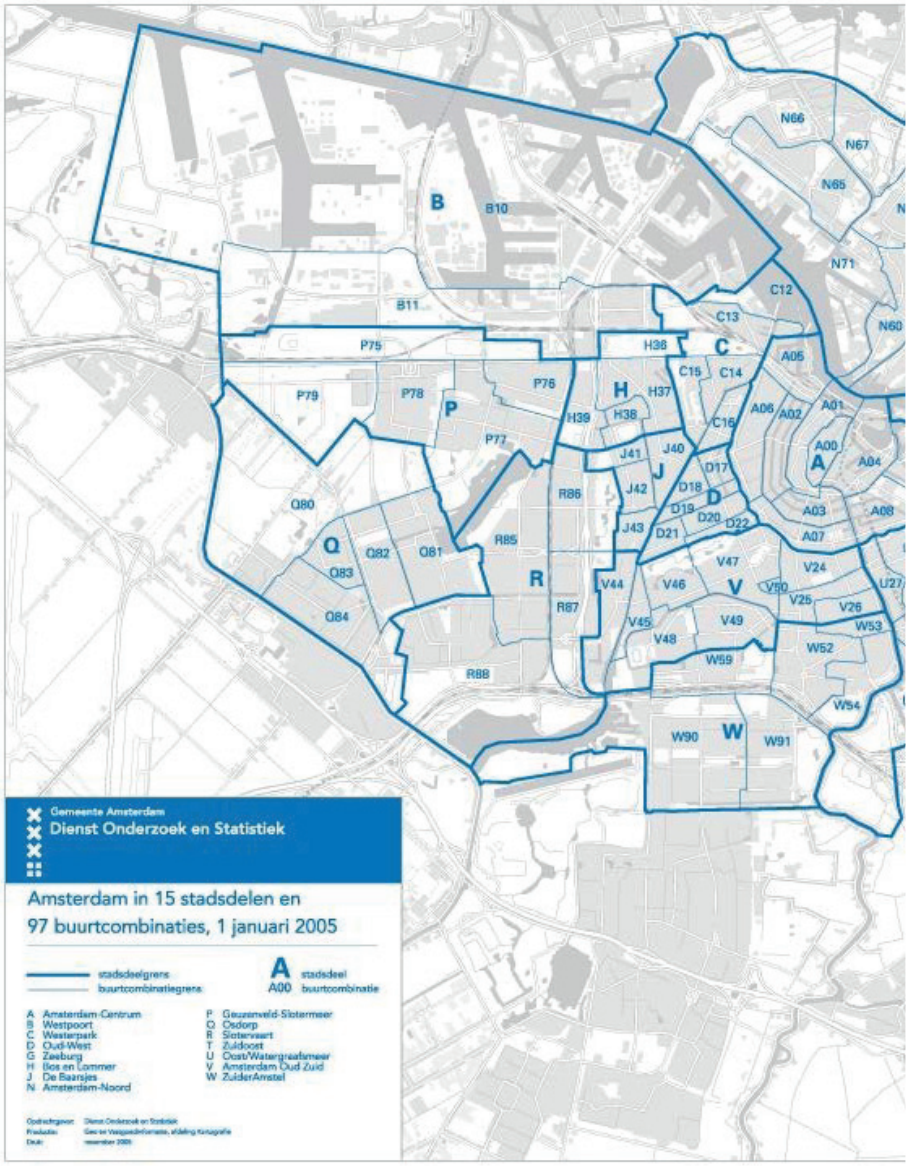
Map 5 Primary schools in Nijmegen in 2003



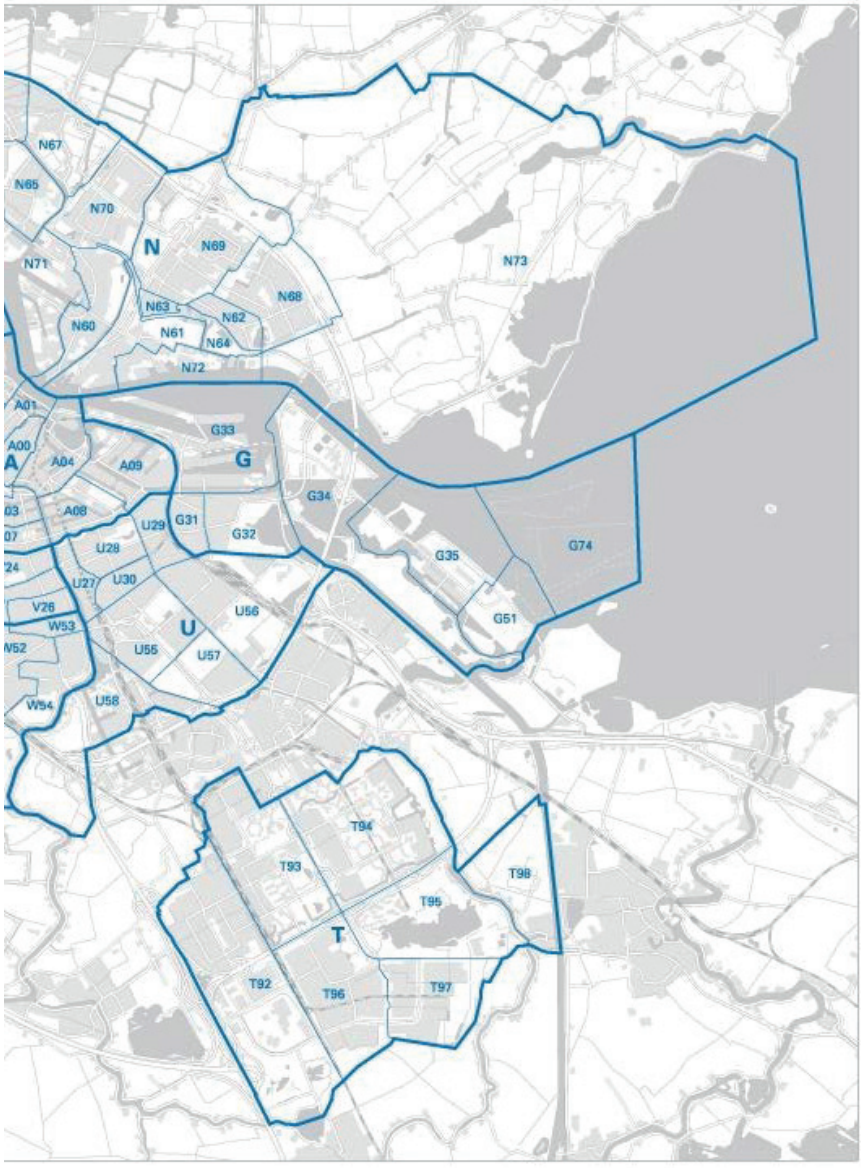
Map 1 *Stadsregio Amsterdam* 'City Region of Amsterdam, i.e. the metropolitan region of Amsterdam' with its 16 neighboring municipalities; city of Amsterdam in red and city of Aalsmeer in blue. Source: http://www.stadsregioamsterdam.nl/ (2014), adapted by LvM

Table 1 Key to Map 2a and 2b of relevant *stadsdelen* and *buurtcombinaties*

Map 2a (Left part of Amsterdam)		*Map 2b (Right part of Amsterdam)*	
Stadsdeel Osdorp	Q	**Stadsdeel** Oost/Watergraafsmeer	U
Osdorp-Midden	Q82	Dapperbuurt	U29
Osdorp-Oost	Q81	Frankendael	U55
		Oosterparkbuurt	U28
Stadsdeel ZuiderAmstel	W	Transvaalbuurt	U30
Scheldebuurt	W52	Weesperzijde	U27
		Stadsdeel Zuidoost	T
		Bijlmer Centrum	T93
		Bijlmer Oost	T94



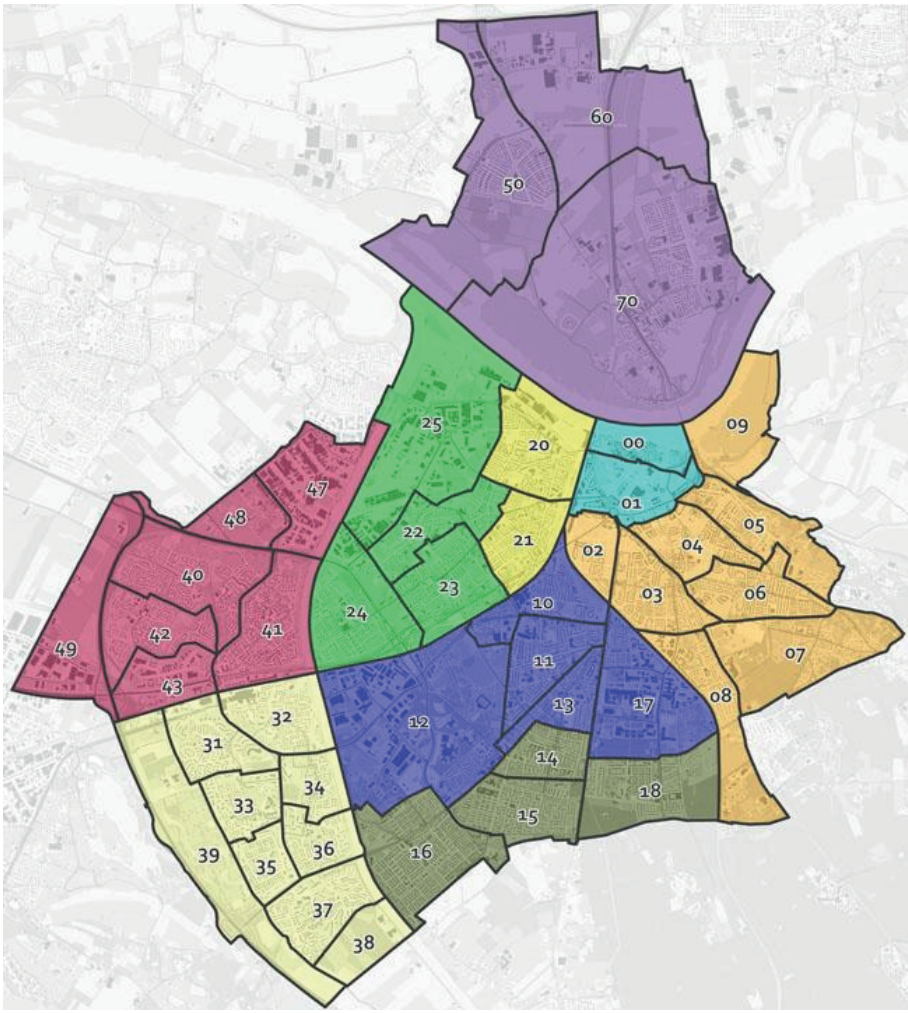
Map 2a Amsterdam with its 15 *stadsdelen* 'parts of town' and 97 *buurtcombinaties* 'combinations of neighborhoods', 1998 till May 1st 2009 (Source: O+S Amsterdam 2005), left part



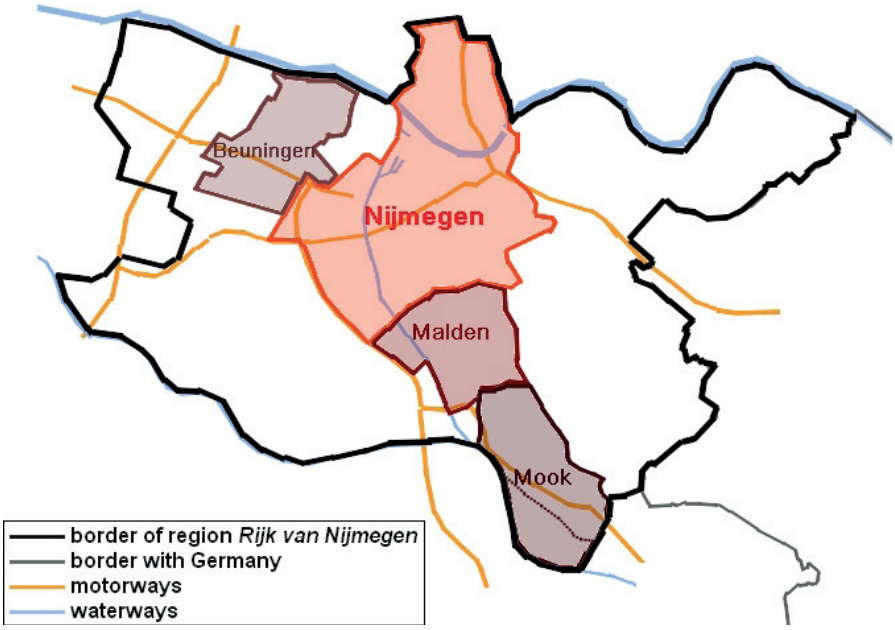
Map 2b Amsterdam with its 15 *stadsdelen* 'parts of town' and 97 *buurtcombinaties* 'combinations of neighborhoods', 1998 till May 1st 2009 (Source: O+S Amsterdam 2005), right part

Table 2 Key to Map 3 of relevant *stadsdelen* and *wijken*

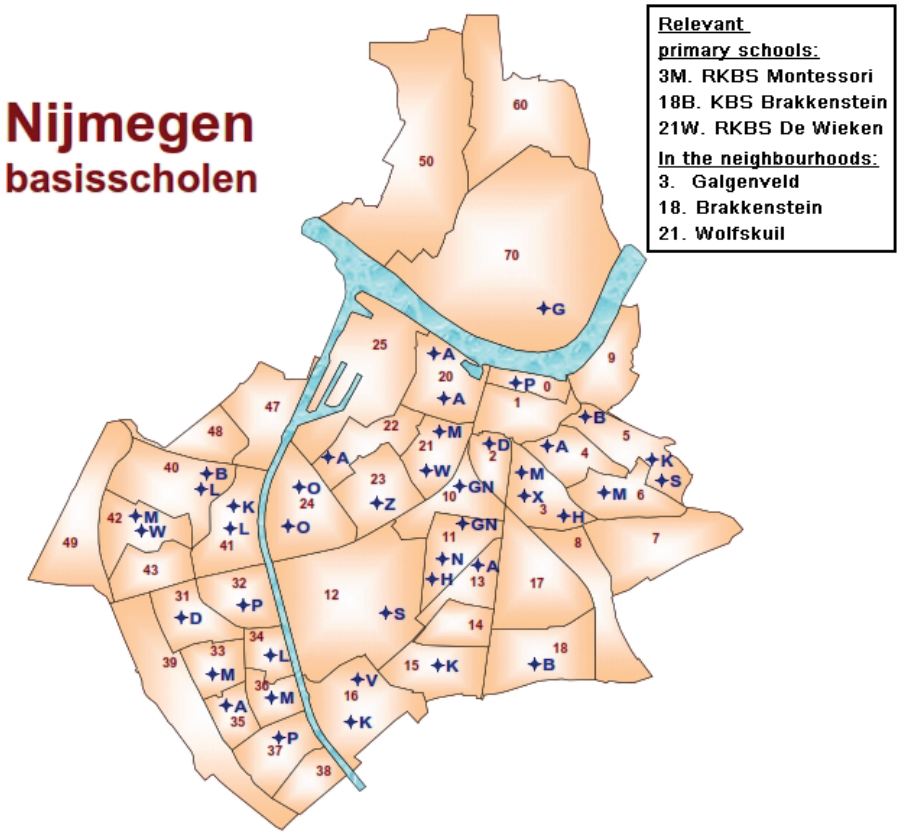
Stadsdeel Oud-West	Bright yellow
Wolfskuil	21
Stadsdeel Nieuw-West	Green
Heseveld	23
Neerbosch-Oost	24
Stadsdeel Midden	Blue
Goffert	12
Hazenkamp	11
Heijendaal	17
Nije Veld	10
St. Anna	13
Stadsdeel Zuid	Gray
Brakkenstein	18
Hatert	16
Stadsdeel Dukenburg	Pale yellow
't Acker	37
Stadsdeel Lindenholt	Pink
Weezenhof	40
Stadsdeel Centrum	Aqua-marine
Beneden-stad	00
Stadsdeel Oost	Orange
Galgenveld	03



Map 3 Nijmegen with its 9 *stadsdelen* 'parts of town' and 44 *wijken* 'neighborhoods' (Source: O&S Nijmegen 2014)



Map 4 Rijk van Nijmegen, lit. 'Land of Nijmegen' with the four municipalities and villages relevant to our study; Created by Linda van Meel, based on the municipalities mentioned by Schulte (1982) and VVV Arnhem Nijmegen (2014)



Map 5 Primary schools in Nijmegen in school year 2001-2002 (O&S Nijmegen 2003). Text box added by LvM

Appendix 4:
Demographic tables of schools and living areas

Table 1 Composition of school population in NIJMEGEN in school year 2001-2002 of the primary schools in our study (Commissie Spreiding en Bereikbaarheid Nijmegen 2003; O&S Nijmegen 2003) (1a)

School	*Wijk*	School year	School population (2)	% non-Western *allochtonen* (3) in base generation (4)	% non-Western *allochtonen* in school population	'Black' school ? (5)	Background recruited core speakers
RKBS De Wieken	Wolfskuil	2001 - 2002	190	46	62	Yes	N12D, N12T, N12M
KBS Brakkenstein	Brakkenstein	2001 - 2002	287	4	4	No	N12C
RKBS Montessori	Galgenveld	2001 - 2002	263	7	9	No	N12C

Table 2 Composition of school population in AMSTERDAM in school year 2006-2007 of the primary schools with core participants in our study (O+S Amsterdam 2008) (1b)

School	*Buurtcombinatie*	School year	% non-Western *allochtonen* (3) in *buurtcombinatie* (4b)	% non-Western *allochtonen* in school	'Black' school ? (5)	Background recruited core speakers
Kunstmagneetschool De Kraal	Transvaalbuurt	2006 - 2007	75,1	90,6	Yes	A12D, A12T, A12M
OBS De Kaap	Transvaalbuurt	2006 - 2007	75,1	97,0	Yes	A12M
Buitenveldertse Montessorischool	Buitenveldert-West	2006 - 2007	31,5	14,6	No	A12C

(1a) The study of commissie Spreiding en Bereikbaarheid Nijmegen (2003) only represents pupils who live ánd attend primary school in Nijmegen. I.e. those who live in, for example, Beuningen and attend school in Nijmegen as well as those who live in Nijmegen and attend school in Beuningen are not included.

(1b) The study of O+S Amsterdam represents all pupils, including those who live outside Amsterdam and attend school in Amsterdam and those of school age who live in Amsterdam but attend school elsewhere.

(2) *School population:* "the number of children attending a school" (Commissie Spreiding en Bereikbaarheid Nijmegen 2003:13, our translation).

(3) *Non-Western allochtonen:* if a child or one of the parents was born in a non-Western country, the child was counted as a non-Western *allochtoon*. Non-Western children are mainly Turkish, Moroccan, Antillean, Aruban and Surinam children, but also (among others) Iranian/Iraqi and Indonesian children. The largest groups are the Turkish and Moroccan children (Commissie Spreiding en Bereikbaarheid Nijmegen 2003:13).

(4a) *Base Generation:* "the number of children from 4 until and including 11 years old and 30% of the 12 year olds, living in the neighborhood of a school" (Commissie Spreiding en Bereikbaarheid Nijmegen 2003:13, our translation).

(4b) The percentage of 'non-Western *allochtonen* in *buurtcombinatie'* concerns the percentage non-Western children of school age living in the 'neighborhood combination'.

(5) *'Black' school:* A school with a majority of *non-Western allochtonen*.

Table 3 Areas of the schools in AMSTERDAM and population by *herkomstgroepering* 'group of origin', displaying the number and percentage of 'white' Dutch (i.e. *autochtoon* 'autochthonous'), people of Turkish and of Moroccan origin of the year of recording for 2007 and of 2006 for the recordings from 2005[1] (O+S Amsterdam 2006, 2007)

background core participants (*n*)	School	*buurtcombinatie / stadsdeel*	Year of recording	population, number of residents, *n* (%)	origin: *autochtoon* 'autochthonous', *n* (%)	origin: Turkey, *n* (%)	origin: Morocco, *n* (%)
A12C (3)	Buitenveldertse Montessorischool	Buitenveldert-West	2007	11,785 (100)	7,570 (64.23)	107 (0.91)	196 (1.66)
		Stadsdeel Zuider-Amstel	2007	46,784 (100)	30,970 (66.20)	505 (1.08)	1,117 (2.39)
A12M (1)	OBS De Kaap	Transvaalbuurt	2007	9,413 (100)	3,312 (35.19)	916 (9.73)	1,795 (19.07)
		Stadsdeel Oost/Watergraafsmeer	2007	58,798 (100)	32,760 (55.72)	2,762 (4.70)	5,549 (9.44)
A12M (2), A12T (3), A12D (2)	Kunstmagneetschool De Kraal	Transvaalbuurt	2005	9,606 (100)	3,341 (34.78)	970 (10.10)	1.808 (18.82)
		Stadsdeel Oost/Watergraafsmeer	2005	58,628 (100)	32,378 (55.23)	2,887 (4.92)	5.676 (9.68)
A20M (1), A20T (1)	ROC ASA Amsterdam Oost Economie	Weesperzijde	2005	4,414 (100)	2,633 (59.65)	69 (1.56)	287 (6.50)
		Stadsdeel Oost/Watergraafsmeer	2005	58,628 (100)	32,378 (55.23)	2,887 (4.92)	5.676 (9.68)

[1] As the classification *herkomstgroepering* was introduced in 2006, the publications from 2005 by O+S Amsterdam do not contain statistical information on this classification. Therefore, we used the number of 2006 in Table 3.

Table 4 Areas of the schools in NIJMEGEN and population by *herkomstgroepering* 'group of origin', displaying the number and percentage of 'white' Dutch (i.e. *autochtoon* 'autochthonous'), people of Turkish and of Moroccan origin of the year(s) of recording (O&S Nijmegen 2014)

background core participants *(n)*	School	*wijk / stadsdeel*	Year(s) of recording	population, number of residents, *n* (%)	origin: *autochtoon* 'autochthonous', *n* (%)	origin: Turkey, *n* (%)	origin: Morocco, *n* (%)
N12C (3)	KBS Brakkenstein	Brakkenstein	2005	3,828 (100)	3,237 (84.56)	23 (0.60)	8 (0.21)
		Stadsdeel Zuid	2005	22,795 (100)	17,033 (74.72)	1,021 (4.48)	433 (1.90)
N12M (3), N12T (3), N12D (3)	RKBS De Wieken	Wolfskuil	2005	5,981 (100)	3,933 (65.76)	784 (13.11)	200 (3.34)
		Stadsdeel Oud-West	2005	12,766 (100)	9,143 (71.62)	993 (7.78)	357 (2.80)
N12C (1)	RKBS Montessori	Galgenveld	2006	6,398 (100)	5,306 (82.93)	14 (0.22)	20 (0.31)
		Stadsdeel Oost	2006	32,615 (100)	26,578 (81.49)	464 (1.42)	236 (0.72)
N20M (1), N20T (1), N20D (1)	ROC Economie (location Marterstraat)	Hazenkamp	2006	5,218 (100)	4,555 (87.29)	17 (0.33)	10 (0.19)
			2007	5,195 (100)	4,542 (87.43)	20 (0.38)	9 (0.17)
		Stadsdeel Midden	2006	17,926 (100)	14,162 (79.00)	665 (3.71)	327 (1.82)
			2007	18,113 (100)	14,353 (79.24)	668 (3.69)	326 (1.80)
N20M (2), N20T (2), N20D (2)	ROC Handel (location Goffertweg)	Goffert	2006	2,687 (100)	2,105 (78.34)	69 (2.57)	27 (1.00)
			2007	2,641 (100)	2,060 (78.00)	72 (2.73)	26 (0.98)
		Stadsdeel Midden	2006	17,926 (100)	14,162 (79.00)	665 (3.71)	327 (1.82)
			2007	18,113 (100)	14,353 (79.24)	668 (3.69)	326 (1.80)
N20C (4)	ROC Zorg en Welzijn (location Vossenlaan)	Hazenkamp	2007	5,195 (100)	4,542 (87.43)	20 (0.38)	9 (0.17)
		Stadsdeel Midden	2007	18,113 (100)	14,353 (79.24)	668 (3.69)	326 (1.80)

Table 5 Reported living areas by participants in AMSTERDAM and population by *herkomstgroepering* 'group of origin', displaying the number and percentage of 'white' Dutch (i.e. *autochtoon* 'autochthonous'), people of Turkish and of Moroccan origin of 2006 (O+S Amsterdam 2006)

Background reporting core participants *(n)*	Reported (recent + previous) living areas, i.e. *buurtcombinatie* and *stadsdeel* (2006)	population, number of residents, *n* (%)	origin: *autochtoon* 'autochthonous', *n* (%)	origin: Turkey, *n* (%)	origin: Morocco, *n* (%)
A12D (2), A12M (3), A12T (2), A20M (1), A20T06 (1).	Stadsdeel Oost/Watergraafsmeer	58,628 (100)	32,378 (55.23)	2,887 (4.92)	5,676 (9.68)
A12D (1)	Frankendael	9,326 (100)	6,018 (64.53)	244 (2.62)	757 (8.12)
A12D (1), A12M (3), A12T (3).	Transvaalbuurt	9,606 (100)	3,341 (34.78)	970 (10.10)	1,808 (18.82)
A20M (1)	Stadsdeel Osdorp	45,465 (100)	21,750 (47.48)	4,329 (9.52)	7,163 (15.75)
A20D (1), A20T (1)	Stadsdeel ZuiderAmstel	46,885 (100)	31,069 (66.27)	533 (1.14)	1,133 (2.42)
A20D (1), A20T (1)	Scheldebuurt	13,176 (100)	9,483 (71.97)	81 (0.61)	180 (1.37)
A20M (1)	Stadsdeel Zuidoost	78,907 (100)	22,946 (29.08)	744 (0.94)	1,511 (1.91)
A20M (1)	Bijlmer Centrum	21,313 (100)	2,831 (13.28)	408 (1.91)	605 (2.48)
A20M (1)	Bijlmer Oost	22,739 (100)	5,093 (22.40)	139 (0.61)	401 (1.76)

Table 6 Reported living areas by participants in NIJMEGEN and population by *herkomstgroepering* 'group of origin', displaying the number and percentage of 'white' Dutch (i.e. *autochtoon* 'autochthonous'), people of Turkish and of Moroccan origin of 2006 (O&S Nijmegen 2014)

background reporting core participants *(n)*	Reported (recent + previous) living areas, i.e. *wijken* and *stadsdelen*	population, number of residents, *n* (%)	origin: *autochtoon* 'autochthonous', *n* (%)	origin: Turkey, *n* (%)	origin: Morocco, *n* (%)
N20C (1), N20M (1)	Stadsdeel Centrum	9,286 (100)	7,547 (81.27)	72 (0.78)	53 (0.57)
N20C (1)	Benedenstad	2,793 (100)	2,222 (79.56)	16 (0.57)	26 (0.93)
N20D (1), N20M (1)	Stadsdeel Dukenburg	23,361 (100)	15,983 (68.42)	674 (2.89)	812 (3.48)
N20D (1)	't Acker	5,673 (100)	4,043 (71.27)	114 (2.01)	113 (1.99)
N20C (1), N20D (1)	Stadsdeel Lindenholt	15,914 (100)	11,585 (72.80)	281 (1.77)	276 (1.73)
N20C (1), N20D (1)	Weezenhof	3,711 (100)	2,978 (80.25)	7 (0.19)	10 (0.27)
N12C (3), N12D (1), N12M (2), N20C (1), N20M (1)	Stadsdeel Midden	17,926 (100)	14,162 (79.00)	665 (3.71)	327 (1.82)
N12C (1), N12D (1), N20M (1)	Goffert	2,687 (100)	2,105 (78.34)	69 (2.57)	27 (1.00)
N12C (2)	Heijendaal	1,748 (100)	1,445 (82.67)	12 (0.69)	0 (0.00)
N12M (2), N20C (1)	Nije Veld	5,125 (100)	3,334 (65.05)	551 (10.75)	283 (5.52)
N12C (1)	St. Anna	3,148 (100)	2,723 (86.50)	16 (0.51)	6 (0.19)
N20M (2)	Stadsdeel Nieuw-West	16,399 (100)	11,896 (72.81)	759 (4.65)	751 (4.60)
N20M (1)	Heseveld	6,071 (100)	4,469 (73.61)	221 (3.64)	307 (5.06)
N20M (2)	Neerbosch-Oost	7,341 (100)	4,921 (67.03)	513 (6.99)	429 (5.84)
N12D (2), N12M (2), N12T (3), N20C (1), N20T (1)	Stadsdeel Oud-West	13,073 (100)	9,411 (71.99)	994 (7.60)	362 (2.77)
N12D (2), N12M (2), N12T (3), N20C (1), N20T (1)	Wolfskuil	6,265 (100)	4,135 (66.00)	789 (12.59)	227 (3.62)
N20C (2)	Stadsdeel Oost	32,615 (100)	26,578 (81.49)	464 (1.42)	236 (0.72)
N12C (2), N20C (1), N20D (1), N20T (2)	Stadsdeel Zuid	22,650 (100)	16,949 (74.83)	1,014 (4.48)	422 (1.86)
N12C (2), N20C (1)	Brakkenstein	3,828 (100)	3,247 (84.82)	31 (0.81)	10 (0.26)
N20T (2)	Hatert	9,607 (100)	6,469 (67.34)	626 (6.52)	299 (3.11)

Table 7 Population of 3 municipalities by *herkomstgroepering* 'group of origin', displaying the number and percentage of 'white' Dutch (i.e. *autochtoon* 'autochthonous'), people of Turkish and of Moroccan origin of the year of recording (CBS 2014)

Background participant (n)	Municipality	Year of recording	population, number of residents, *n* (%)	origin: *autochtoon* 'autochthonous', *n* (%)	origin: Turkey, *n* (%)	origin: Morocco, *n* (%)
N12D (1)	Beuningen	2005	25,291 (100)	22,520 (89.04)	42 (0.17)	96 (0.38)
N20C (1)	Heumen	2007	16,673 (100)	14,758 (88.51)	22 (0.13)	20 (0.12)
	Mook en Middelaar	2007	8,029 (100)	7,017 (87.40)	3 (0.04)	9 (0.11)

Contributions

Chapters 2, 3, 4 and 5 of this dissertation are based on scientific manuscripts with the PhD candidate as first author. Supervisors Frans Hinskens and Roeland van Hout featured as co-authors. Below the contributions of the PhD candidate and the co-authors are described to allow for a full assessment of the candidate's work.

Research contributions for Chapters 2, 3, 4 and 5

PhD candidate

The candidate set up and implemented all data extraction from the recordings and coded and analyzed the data.

Supervisors

The supervisors gave frequent feedback and supervision.

Frans Hinskens selected the phenomena to be researched, the corresponding research questions and the analyses of the internal conditioning. Roeland van Hout advised on the coding and exploration of the data and the statistical analyses.

Phonetic analyses and part of the coding

Sander van der Harst performed various phonetic analyses for Chapters 2, 3 and 4. He also coded a quarter of the /a:/ and /ɑ/ data (Chapter 4).

Data by others in Chapter 5 (the co-variation study)

The data of *hun* ('them', subject 3 pl.) was extracted and coded by several interns under supervision of the PhD candidate. The data of neuter gender marking was extracted and coded by Ariën van Wijngaarden.

Writing contributions for Chapters 2, 3, 4 and 5

The papers were written by the PhD candidate in collaboration with the two supervisors. All three contributed to all sections. The PhD candidate wrote most of the method and results sections. Frans Hinskens contributed to the introduction, conclusion and discussion sections. Roeland van Hout contributed to the method, results, conclusion and discussion sections.

Samenvatting (Dutch summary)

Taalvariatie kan zowel voortkomen uit taalcontact als uit tendensen die inherent zijn aan taal. Over variatie als gevolg van taalcontact weten we over het algemeen waar en wanneer deze ontstaat, in tegenstelling tot inherente variatie, een bron van taalvariatie die altijd aanwezig is. Deze twee bronnen van heterogeniteit hebben geleid tot twee takken van taalkundig onderzoek die zich min of meer apart van elkaar hebben ontwikkeld. De ene tak, taalcontactonderzoek, richt zich op codewisselingen en ontleningen, alsook op de studie van pidgins, creooltalen en gemengde talen. De andere tak, kwantitatief sociolinguïstisch onderzoek, onderzoekt verschillen in taalgemeenschappen en verschillen tussen sociolecten en dialecten.

Er bestaat ook een type taalvariatie waarbij deze twee takken samenkomen: etnolectische variatie. Dit is een recent thema binnen het onderzoek naar tweetaligheid, taalcontact en taalvariatie. Omdat etnolectische variatie geleidelijk en continu van aard is, met variatie tussen de standaardtaal en de eigen etnolectische vorm ervan, zijn er voor het onderzoek benaderingen vanuit de kwantitatieve sociolinguïstiek nodig.

Dit proefschrift maakt deel uit van het project *'The Roots of Ethnolects, an Experimental Comparative Study'* [De oorsprong van etnolecten, een experimentele vergelijkende studie]. Het project is opgezet met als doel de wortels, d.w.z. de verschillende bronnen, van etnolect-variatie te onderzoeken. Het project concentreert zich op de opkomst, positie en sociale verspreiding van twee jonge etnolecten van het Nederlands in Amsterdam en Nijmegen, en is gericht op variabele fonologische en morfo-syntactische verschijnselen, te bestuderen bij jonge, tweede-generatie immigranten van Marokkaanse en Turkse afkomst.

Voor het project werd een gestructureerd design opgezet. De informanten waren 10-12 jaar en 18-20 jaar oude mannelijke inwoners uit Amsterdam en Nijmegen met Marokkaans-Nederlandse, Turks-Nederlandse en 'witte' Nederlandse achtergronden. Bij de 'witte', autochtone Nederlandse jongeren is onderscheid gemaakt tussen sprekers met en zonder interetnisch netwerk. De informanten zijn geboren en getogen in Amsterdam dan wel Nijmegen. Het is onmogelijk om politiek correcte labels aan de verschillende groepen te geven. In dit proefschrift is gekozen voor *Turks-Nederlands*, *Marokkaans-Nederlands* en *'wit' Nederlands*.

Binnen het project zijn zowel *spontane spraakdata* als *gecontroleerde spraakdata* verzameld. In dit proefschrift wordt echter alleen gebruik gemaakt van de spontane spraakdata. Voor de spontane spreekdata zijn opnames gemaakt van gesprekken tussen telkens twee jongeren, waarbij het streven was dat iedere jongere

deelnam aan drie tot vier gesprekken van ongeveer een uur. De gesprekspartner was ofwel een jongere met dezelfde achtergrond (een ingroep-gesprek) ofwel een jongere met een andere achtergrond (een uitgroep-gesprek), met de bedoeling dat elke jongere met leeftijdsgenoten van alle drie de achtergronden sprak. De 'witte' Nederlandse jongeren zonder interetnisch netwerk voerden alleen gesprekken met andere autochtone Nederlandse jongeren zonder interetnisch netwerk. Ze dienden als controlegroep voor vergelijking met de ingroep-gesprekken van de 'witte' Nederlandse jongeren met een interetnisch netwerk. Om een zo natuurlijk mogelijke conversatie te garanderen was de interviewer indien mogelijk alleen aan het begin en aan het einde van het gesprek aanwezig.

Voor het onderzoek in dit proefschrift is een 'kerncorpus' van 51 sprekers uit de database geselecteerd die ook daadwerkelijk gesprekken met alle drie de achtergronden hebben gevoerd. De sprekers zijn zo veel mogelijk gelijkmatig verdeeld met betrekking tot hun achtergrond (Marokkaans-Nederlands, Turks-Nederlands en 'wit' Nederlands), leeftijd (10–12 versus 18–20 jaar), woonplaats (Amsterdam versus Nijmegen) en, in geval van een wit Nederlandse achtergrond, ook de aan- of afwezigheid van interetnische contacten met Marokkaans-Nederlandse en Turks-Nederlandse achtergronden. Een schematisch overzicht is te vinden in Tabel 4 van Hoofdstuk 1.

Het onderzoek in dit proefschrift richt zich op zeven onderzoeksvragen. Vier van deze vragen richten zich op de herkomst (*roots*) van variatiepatronen: (1) In welke mate zijn variatiepatronen gebaseerd op interferentie van de *heritage*-talen van de sprekers met Marokkaanse en Turkse achtergronden? (substraat-effecten), (2) In welke mate kunnen variatiepatronen gerelateerd worden aan de invloed van lokale niet-standaard variëteiten, d.w.z. de dialectverschillen tussen Amsterdam en Nijmegen? (regionale effecten), (3) In welke mate zijn de patronen en regels van het endogeen Nederlands (zoals bv. progressieve stemassimilatie) van toepassing op de variatiepatronen binnen sprekers met Marokkaanse en Turkse achtergronden? (structurele effecten), (4) In welke mate kunnen we variatiepatronen relateren aan kenmerken die het resultaat zijn van processen van tweedetaalverwerving? (tweedetaalverwervingseffecten)

De vijfde onderzoeksvraag richt zich op stijlverschuivingen (*style-shifting*): (5) In welke mate kunnen sprekers schakelen tussen verschillende variatiepatronen, afhankelijk van de achtergrond van hun gesprekspartner?

De afsluitende twee onderzoeksvragen gaan over de sociale verspreiding van de variatiepatronen: (6) Verschillen de sprekers met Marokkaanse en Turkse achtergronden in hun variatiepatronen of delen ze een meer globale '*non-native'* identiteit? (interetnische convergentie-effecten) (7) Zijn er aanwijzingen voor de

verspreiding van etnische patronen naar 'witte' Nederlandse jongeren? (*native* convergentie-effecten)

Om de onderzoeksvragen te kunnen beantwoorden zijn vier Nederlandse fonemen onderzocht: /z/, /ɛi/, /a:/ en /ɑ/, alsook het contrast tussen de laatste twee fonemen. De fonemen en de variatie daarin hebben een drievoudige oorsprong: de *heritage*-talen van de Marokkaans- en Turks-Nederlandse sprekers, de dialecten van Amsterdam en Nijmegen, en tweedetaalverwervings-verschijnselen.

Een van de kenmerken die opvallen bij de onderzochte Marokkaans-Nederlandse en Turks-Nederlandse sprekers in beide steden is een karakteristieke dentale stemhebbende uitspraak van /z/. Deze dentale realisatie heeft zijn oorsprong in de verschillende Marokkaanse talen en komt niet voor in endogeen Nederlands. In Hoofdstuk 2 is deze niet-standaard Nederlandse dentale realisatie van /z/ onderzocht in woordinitiële positie. Ook is er gekeken in hoeverre de Nederlandse verstemlozingsregels toegepast worden door de verschillende groepen. Hiervoor is gekeken naar /z/ in woordinitiële positie volgend op obstruenten, sonoranten of vocalen.

Hoofdstuk 3 onderzoekt de diftong (tweeklank) /ɛi/. Dit foneem bestaat niet als zodanig in de Marokkaanse en Turkse talen. In de Nederlandse taal is er veel regionale en sociale uitspraakvariatie. In het dialect van Amsterdam verschilt de uitspraak van die van Nijmegen. De belangrijkste vraag in dit deelonderzoek is of /ɛi/ betrokken is bij processen van sociale (geografische en etnische) herdistributie. Hiertoe zijn twee variabelen onderzocht: monoftongering en de articulatieplaats van het eerste, prominente element van /ɛi/.

In Hoofdstuk 4 worden de fonemen /a:/ en /ɑ/ onderzocht, met een focus op het contrast tussen beiden. Het lang-kort of gespannen-ongespannen contrast is essentieel in het Nederlands. Er wordt gekeken hoe sprekers van *syllable-timed* talen zoals Turks, Arabisch en Berber, die tweetalig opgroeien met een extra *stress-timed* taal, in dit geval Nederlands, omgaan met dit contrast dat waarschijnlijk niet stelselmatig in het T2-Nederlands van hun ouders voorkomt. Beide fonemen en hun contrast zijn onderzocht voor duur en plaats van articulatie, in positie voorafgaand aan obstruenten en aan nasalen.

Het feit dat de linguïstische variabelen in hoofdstukken 2, 3 en 4 verschillende oorsprongen hebben, roept de vraag op of er samenhangen zijn tussen het gebruik van de kenmerken die een etnolect vormen, dus of de verschijnselen coherent zijn. Hoofdstuk 5 onderzoekt op welke manier de linguïstische variabelen aan elkaar gekoppeld kunnen worden en of de coherentiepatronen verklaard kunnen worden door buitentalige (sociale) factoren, talige factoren of door beide. Daarvoor is de co-variatie tussen de verschillende linguïstische variabelen onderzocht, d.w.z. niet-

toevallige correlaties tussen twee of meer linguïstische variabelen in het taalgebruik van leden van een specifieke (geografische, sociale of culturele) groep.

Wat zijn de antwoorden op de zeven onderzoeksvragen?

1. Substraat-effecten

In welke mate zijn variatiepatronen gebaseerd op interferentie van de heritage*-talen van de sprekers met Marokkaanse en Turkse achtergronden?*

Zowel de stemhebbendheid van /z/ na obstruenten als de dentale realisatie van /z/ zijn exogene, 'niet-Nederlandse' verschijnselen. De stemhebbende uitspraak van /z/ na een obstruent (als in *to[dz]iens* of *o[bz]oek*) heeft zeker geen Nederlandse oorsprong. De dentale realisatie van /z/ heeft zijn oorsprong in de talen van de Marokkaans-Nederlandse sprekers en is overgenomen door de Turks-Nederlandse sprekers.

2. Regionale effecten

In welke mate kunnen variatiepatronen gerelateerd worden aan de invloed van lokale niet-standaard variëteiten, d.w.z. de dialectverschillen tussen Amsterdam en Nijmegen?

Amsterdamse sprekers lieten meer verstemlozing van /z/ na een sonorant of vocaal zien dan de Nijmeegse sprekers, vooral in het geval van de 18-20-jarigen. Ook gebruikten de Amsterdamse sprekers meer open eerste elementen van de diftong /ɛi/ dan de Nijmeegse sprekers. Daarnaast hebben zij een langere /ɑ/ voor nasalen, een meer naar voren uitgesproken /ɑ/ (in zowel obstruent als nasaal contexten) en een meer naar achteren uitgesproken /a:/ in gesloten lettergrepen (in beide contexten) dan de Nijmeegse sprekers. Dit komt overeen met de stadsdialecten.

Nijmeegse sprekers hebben een meer open uitspraak van /ɛ/ in de frase *heb je/ik*, een verschijnsel dat voor het co-variatie-onderzoek onder de loep genomen is. De hoogte van /ɛ/ vormt, samen met drie andere verschijnselen die fonetische variatie in de plaats van articulatie (hoog/laag, voor/achter) vertonen, een cluster van linguïstische variabelen die verschillende uitspraken hebben in Amsterdamse en Nijmeegse sociolecten. Er was een duidelijk onderscheid tussen de Amsterdamse en Nijmeegse sprekers in onze data voor dit cluster.

3. Structurele effecten

In welke mate zijn de patronen en regels van het endogeen Nederlands (zoals bv. progressieve stemassimilatie) van toepassing op de variatiepatronen binnen sprekers met Marokkaanse en Turkse achtergronden?

Deze vraag is in het bijzonder relevant voor de verstemlozing van /z/. In alle

variëteiten van het Nederlands wordt de /z/ stemloos na een obstruent. De verstemlozing van /z/ volgend op obstruenten gebeurt beduidend minder frequent bij de Turks- en Marokkaans-Nederlandse sprekers.

Voor de diftong /ɛi/ blijkt dat monoftongering afhankelijk is van het etymologische verschil tussen <ei> en <ij> (dat nog steeds fonologisch is in verschillende groepen Nederlandse dialecten), de positie ten opzichte van de syllabegrens (d.w.z. open vs. gesloten syllabes) en de frequentie van voorkomen van lemmata. Echter, op een meer gedetailleerd niveau blijken de resultaten moeilijk te interpreteren. Het corpus aan spraakdata in dit onderzoek is helaas te klein om deze interessante linguïstische conditionering in meer detail te onderzoeken.

4. Tweedetaalverwervingseffecten

In welke mate kunnen we variatiepatronen relateren aan kenmerken die het resultaat zijn van processen van tweedetaalverwerving?

Deze vraag is van toepassing op de afstanden tussen /a:/ en /ɑ/ waarbij we invloed verwachtten van tweedetaalverwerving als gevolg van het typologisch onderscheid tussen *syllable-* en *stress-timed* systemen, vertegenwoordigd door het Nederlands aan de ene kant en de *heritage*-talen aan de andere kant. In tegenstelling tot het Nederlands kennen de *heritage*-talen geen fonologisch contrast tussen ongespannen en gespannen klinkers. De Turks- en Marokkaans-Nederlandse sprekers blijken echter hetzelfde lengte-onderscheid te maken als hun 'witte' leeftijdsgenoten. Het verwachtte tweedetaalverwervingseffect kwam dus niet naar voren.

5. Stijlverschuivingseffecten

In welke mate kunnen sprekers schakelen tussen verschillende variatiepatronen, afhankelijk van de achtergrond van hun gesprekspartner?

De achtergrond van de gesprekspartner is in meerdere opzichten van invloed geweest op het gedrag van de sprekers. Deze invloed was het duidelijkst voor de dentalisatie van /z/. Sprekers van de Turks- en Marokkaans-Nederlandse groepen vertoonden de zwakste / minst frequent dentalisatie in hun gesprekken met de autochtone Nederlandse sprekers en de sterkste / meest frequente in gesprekken met sprekers van hun eigen achtergrond. Dentalisatie is voor beide groepen een ingroep-kenmerk. Dit gedrag is zelfs nog duidelijker voor de Marokkaans-Nederlandse sprekers dan voor de Turks-Nederlandse sprekers. Er komen complexe stijleffecten voor in de uitspraakvariatie in de diftong /ɛi/ die gerelateerd zijn aan de achtergrond van de gesprekspartner. De patronen zijn soms moeilijk te duiden omdat er verschillende effecten tegelijkertijd als ook complexe interactie-effecten een rol lijken te spelen. Ook bij de realisatie van de klinkers /ɑ/ en /a:/ (zowel voor duur als

voor plaats van articulatie en vóór zowel obstruenten als nasalen) kwamen diverse stijleffecten naar voren die gerelateerd zijn aan de achtergrond van de gesprekspartner.

Bell's (1984) stelling dat kenmerken van spraak gebruikt in interactie met leden van bepaalde groepen weerspiegeld worden in de spraak van deze groepen, komt het duidelijkst tot uiting in de dentalisatie van /z/. Sprekers gebruiken toenemende hoeveelheden dentalisatie van /z/ wanneer ze spreken met leden van de drie groepen overeenkomstig het patroon 'wit' Nederlands < Turks-Nederlands ≤ Marokkaans-Nederlands, wat de hoeveelheden dentalisatie in de spraak van de drie groepen spiegelt. Voor de overige variabelen, de klinkers, is Bell's stelling minder duidelijk van toepassing. Het komt alleen naar voren in de monoftongering van /ɛi/ bij de 20-jarige Turks-Nederlandse Amsterdammers en in de variatie in de duur van /a:/ voor obstruenten bij de Marokkaans-Nederlandse sprekers. In het laatste geval vertoont het taalgebruik van de Marokkaans-Nederlandse sprekers tegenover hun gesprekspartners hetzelfde patroon als de spraak van de leden van de groepen: Marokkaans-Nederlands ≤ Turks-Nederlands < 'wit' Nederlands.

6. Interetnische convergentie-effecten

Verschillen de sprekers met Marokkaanse en Turkse achtergronden in hun variatiepatronen of delen ze een meer globale 'non-native' *identiteit?*

Voor de onderzochte fonemen is er nauwelijks enig verschil tussen de Turks-Nederlandse en Marokkaans-Nederlandse sprekers. Het lijkt erop dat dentalisatie van /z/, dat een Marokkaanse oorsprong heeft, een markeerder is geworden van een meer globale *'non-native'* identiteit en dat het zich begint te ontwikkelen tot een 'multi-etnolect' kenmerk. De Marokkaans-Nederlandse sprekers gaan echter voorop in het gebruik van gedentaliseerde varianten aangezien ze meer in staat zijn tot aanpassing aan de gesprekspartner dan de Turks-Nederlandse sprekers. De monoftongering van /ɛi/, die vanouds kenmerkend is voor de Amsterdamse en Nijmeegse stadsdialecten en die vooral door leden van de lagere sociale klassen gehanteerd werd, lijkt ge-etnolectiseerd te zijn door zowel de Turks- als Marokkaans-Nederlandse sprekers. De monoftongische variant blijkt verworpen te worden door de autochtone Nederlandse sprekers. Dit verschijnsel ondergaat met andere woorden een sociale herdistributrie.

7. *Native* convergentie-effecten

Zijn er aanwijzingen voor de verspreiding van etnische patronen naar 'witte' Nederlandse jongeren?

De 'witte' Nederlandse sprekers nemen de etnische gedentaliseerde varianten niet over. Deze autochtone Nederlandse sprekers lijken ook vast te houden aan hun

eigen patronen met betrekking tot de stemhebbendheid van /z/ aangezien ze meer verstemlozen dan de Turks- en Marokkaans-Nederlandse sprekers.

Zoals hierboven al beschreven, ondergaat monoftongering van /ɛi/ een sociale herdistributie waarbij de monoftongische variant verworpen wordt door de 'witte' Nederlandse sprekers en ge-etnolectiseerd door de Turks- en Marokkaans-Nederlandse sprekers. Dus voor de verschijnselen onderzocht in dit proefschrift lijkt er geen *crossing* (Rampton 1995) voor te komen. Er is een duidelijke scheidslijn tussen de autochtone Nederlandse sprekers aan de ene kant en de Turks- en Marokkaans-Nederlandse sprekers aan de andere kant.

Curriculum vitae

Linda van Meel was born on October 26, 1984 in Rotterdam, Netherlands. She earned a Bachelor's degree in Linguistics from the Radboud Universiteit Nijmegen in 2006 and completed a Master's degree at the same university in 2008. Her research focused on the acquisition of Dutch by Turkish and Moroccan second language learners.

In 2009, she started as a research assistant in the project 'Roots of Ethnolects' (RoE), working at the Radboud Universiteit Nijmegen as well as at the Meertens Instituut, Amsterdam. In 2010, she took over the PhD position on phonological aspects in the same project, which has resulted in this dissertation. In addition to her research, she designed and maintained the RoE project-website and she took care of the RoE database.

She also participated in the Languages in Contact group (LinC) for which she (a) organized the internal colloquia (until April 1, 2014), (b) maintained both the internal as well as the public LinC website (also until April 1, 2014), and (c) was involved in the initial stage of the curation of the databases within the LinC group by making an inventory of (the status of) all databases.